

Phenomenology of a Pseudo-Scalar Inflaton: Naturally Large Nongaussianity

Neil Barnaby, Ryo Namba, Marco Peloso

*School of Physics and Astronomy, University of Minnesota, Minneapolis, MN 55455,
USA*

E-mail: barnaby@physics.umn.edu, namba@physics.umn.edu,
peloso@physics.umn.edu

ABSTRACT: Many controlled realizations of chaotic inflation employ pseudo-scalar axions. Pseudo-scalars ϕ are naturally coupled to gauge fields through $c\phi F\tilde{F}$. In the presence of this coupling, gauge field quanta are copiously produced by the rolling inflaton. The produced gauge quanta, in turn, source inflaton fluctuations via inverse decay. These new cosmological perturbations add incoherently with the “vacuum” perturbations, and are highly nongaussian. This provides a natural mechanism to generate large nongaussianity in single or multi field slow-roll inflation. The resulting phenomenological signatures are highly distinctive: large nongaussianity of (nearly) equilateral shape, in addition to detectably large values of both the scalar spectral tilt and tensor-to-scalar ratio (both being typical of large field inflation). The WMAP bound on nongaussianity implies that the coupling c of the pseudo-scalar inflaton to any gauge field must be smaller than about $10^2 M_p^{-1}$.

KEYWORDS: Inflation, Nongaussianity, Axions.

Contents

1. Introduction	2
2. Overview of the Mechanism	4
2.1 Production of Gauge Field Fluctuations	5
2.2 Backreaction Effects	6
2.3 Inverse Decay Effects and Inflaton Perturbations	8
3. Correlation Functions	9
3.1 Two-Point Correlation Function	10
3.2 The Primordial Power Spectrum	13
3.3 Three-Point Correlation Function	14
3.4 The Bispectrum and Nonlinearity Parameter	15
3.5 Power spectrum of the Tensor Modes	16
4. Phenomenology of Axion Inflation	18
4.1 COBE Normalization and Spectral Tilt	18
4.2 Tensor-to-Scalar Ratio	20
4.3 Nongaussianity	20
4.3.1 The Size and Running of the Nongaussianity	21
4.3.2 The Shape of the Nongaussianity	22
4.4 Large Field Inflation	25
5. Cosmological Perturbation Theory	26
5.1 ADM Formalism	27
5.2 Quadratic Action	28
5.3 Cubic Action	28
5.4 The Equations of Motion	29
5.5 The Curvature Perturbation	30
6. Models of Axion Inflation	30
6.1 Natural Inflation	31
6.2 Double-Axion Inflation	32
6.3 N-flation	32
6.4 Axion Monodromy Inflation	33
6.5 Axion/4-Form Mixing	34
6.6 Dante's Inferno	34
6.7 Multi-Field Scenarios	34
7. Conclusions	35

1. Introduction

Primordial inflation provides a simple mechanism to resolve the conceptual difficulties of the standard Big Bang cosmology and has enjoyed great phenomenological success in accounting for the properties of the observed Cosmic Microwave Background (CMB) anisotropies. As such, inflation has become the dominant paradigm for the early universe. In spite of this success, however, a compelling particle physics realization is still lacking. The key obstruction is the requirement of a suitably flat scalar potential, $V(\varphi)$. Successful inflation requires $\epsilon, |\eta| \ll 1$ where the slow roll parameters are defined as

$$\epsilon \equiv \frac{M_p^2}{2} \left(\frac{V'}{V} \right)^2, \quad \eta \equiv M_p^2 \frac{V''}{V} \quad (1.1)$$

Here prime denotes derivative with respect to φ , and $M_p \simeq 2.4 \cdot 10^{18}$ GeV is the reduced Planck mass. These parameters are notoriously sensitive to Ultra-Violet (UV) physics. For example, even generic Planck-suppressed corrections to $V(\varphi)$ may contribute $\Delta\eta = \mathcal{O}(1)$, thus spoiling inflation. This UV sensitivity represents a technical fine tuning problem which must be addressed in any particle physics model of inflation.

It is conceivable that dangerous corrections to ϵ, η may be absent as a result of fine-tuning [1], or that the requirement of a flat potential can be evaded by invoking somewhat exotic effects such as dissipation [2, 3], small sound speed [4] or higher derivative corrections [5]. However, perhaps the simplest and most cogent way to realize $\epsilon, |\eta| \ll 1$ in a natural way is by assuming that the inflaton φ is a Pseudo-Nambu-Goldstone-Boson (PNGB) [6, 7, 8, 9, 10, 11, 12, 13, 14, 15]. In this case the inflaton enjoys a shift symmetry $\varphi \rightarrow \varphi + \text{const}$, which is broken either explicitly or by quantum effects. In the limit of exact symmetry we must have $\epsilon = \eta = 0$, thus dangerous contributions to the slow roll parameters are controlled by the smallness of symmetry breaking. Moreover, PNGBs like the axion are ubiquitous in particle physics: they arise whenever an approximate global symmetry is spontaneously broken and are plentiful in string theory compactifications.

The idea of invoking a PNGB to obtain a natural realization of inflation is more than 20 years old. The first model, natural inflation [6], exploited the periodic potential

$$V(\varphi) = \Lambda^4 \left[1 - \cos \left(\frac{\varphi}{f} \right) \right] \quad (1.2)$$

which arises from nonperturbative effects and breaks the continuous shift symmetry down to a discrete subgroup $\varphi \rightarrow \varphi + (2\pi)f$. Unfortunately, this model is compatible with observation only when the axion decay constant is $f > M_p$ [16], a regime that may be impossible to realize in a controlled effective field theory because it suggests a global symmetry broken *above* the quantum gravity scale [8]. It has also been conjectured that $f > M_p$ cannot be realized in a controlled limit of string theory [17]. Fortunately, this difficulty can be easily evaded, for example by considering two [9] or more [10, 11] axion fields, extra dimensions [8], or by exploiting the non-periodic contributions to $V(\varphi)$ that arise from wrapping branes on suitable cycles [12, 13]. Currently, there exist a number of natural, controlled realizations of axion inflation with an axion scale f a few orders of magnitude smaller than M_p but that nevertheless behave effectively as large field inflation models [9, 10, 11, 12, 13, 14, 15].

In any axion inflation model, the inflaton is expected to couple to some gauge field A_μ via interactions of the type

$$\mathcal{L}_{\text{int}} = -\frac{\alpha}{4f} \varphi F^{\mu\nu} \tilde{F}_{\mu\nu} \quad (1.3)$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the field strength and $\tilde{F}^{\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\alpha\beta} F_{\alpha\beta}$ is its dual. The strength of the interaction is controlled by the decay constant, f , and by the dimensionless parameter α . While α is in principle a model dependent quantity, from the perspective of effective field theory we generally expect it to be order unity. On the other hand, Ref. [15] has provided concrete examples (in multi-field or extra-dimensional models) which can result in α greater or much greater than one. Thus, for controlled effective field theory realizations of axion inflation it is very natural to have $\alpha/f \gg M_p^{-1}$, in which case the interaction (1.3) is much stronger than gravitationally suppressed. In this work, we consider in detail the cosmological implications of the interaction (1.3), which has been neglected in nearly all previous studies. Our analysis is quite general: the general logic of effective field theory requires the inclusion of an interaction (1.3) whenever ϕ is pseudo-scalar.

In [15] it was shown that energy dissipation into gauge fields, via the interaction (1.3), can slow the motion of the inflaton on a very steep potential. A more conservative approach was adopted in our recent work [18], where we note that the interaction (1.3) can have a profound impact on the phenomenology of the model, even in the conventional slow roll regime. The underlying physics is as follows. The motion of the inflaton amplifies the fluctuations of the gauge field, δA , which in turn produce inflaton fluctuations via *inverse decay*: $\delta A + \delta A \rightarrow \delta\varphi$. When $f/\alpha \lesssim 10^{-2} M_p$, which is very natural in models that admit a UV completion, this new source of perturbations actually dominates over the usual fluctuations from the vacuum. In this regime, all previous studies of axion inflation are invalid. Our analysis is phenomenological: we use CMB data to place observational limits on the coupling α/f without making any specific assumptions about the microphysical origins of the model. We believe that this study should serve as motivation for a case-by-case analysis of the allowed values of α in various explicit realizations of axion inflation.

In this paper, we reconsider the cosmological fluctuations in axion inflation, extending significantly our previous work [18]. We are motivated, in part, by the recent surge of interest in computing and measuring nongaussian effects in the CMB (see [19] for a recent review). Nongaussian statistics, such as the bispectrum, provide a powerful tool to discriminate between the plethora of inflationary models in the literature and may provide a valuable window into the detailed physics of the very early universe. However, a single decoupled scalar field in slow roll is well known to produce an undetectably small signal [20, 21, 22]. The reason for this is intuitively easy to understand: nongaussianity is a measure of the strength of interactions, while in the vanilla scenario the requirement of a flat potential typically also constrains interactions to be weak. To evade this no-go result and obtain an observably large signal, previous studies have invoked non-standard field theories (with small sound speed [23] or higher derivatives [5]) or initial conditions [23, 24, 25, 26], potentials with sharp features [27, 28], dissipative effects [3], fine-tuned inflationary trajectories [29] or post-inflationary effects (such as preheating [30, 31]). In the work [18], on the other hand, the no-go results of [20, 21, 22] are circumvented in a

very novel way: the interaction that gives rise to large nongaussianity, eqn. (1.3), does not play any role in the background dynamics and is thus unconstrained by the requirement of slow roll.

In terms of the broader picture we believe that this work – along with [18] and the previous studies [32, 33, 34] – suggests that the conventional lore concerning the difficulty of obtaining large nongaussianity may have been excessively conservative. One generically expects that inflaton to couple to some fields which do not play any role in driving inflation. Such interactions are unavoidable from an effective field theory perspective and are (at least to some extent) necessary in order to successfully reheat the universe after inflation. Refs. [18, 32, 33, 34] provide explicit examples demonstrating how the consistent inclusion of such interactions can radically modify the phenomenology of inflation, via particle production effects. This work represents a challenge to the conventional lore that nongaussianity is a “smoking gun” signature of non-standard inflationary dynamics by illustrating explicitly that perhaps the simplest and best-motivated particle physics models of inflation are *already* constrained by existing observational limits on nongaussianity.

This paper is organized as follows. In section 2 we provide an overview of the mechanism and briefly summarize our previous work [18]. In section 3 we compute the correlation functions of the curvature fluctuations and tensor (gravity wave, GW) perturbations. Section 4 studies the resulting phenomenology. We have written this section in self-contained way, so that a reader who is interested only in the observational predictions can skip the previous more technical section. In section 5 we perform a complete computation of the perturbations, that includes also the metric perturbations. We show that metric perturbations can be neglected (at leading order), in the computation of the density and GW correlators sourced by the gauge field. In section 6 we review most of the existing models of axion inflation, and briefly discuss the implications of our findings for these models. Finally, in section 7, we conclude.

2. Overview of the Mechanism

In this section we provide a brief overview of the production of gauge field fluctuations and the subsequent inverse decay effects in axion inflation. This section is largely review of [18].

We consider a simple theory of a PNGB inflaton interacting with a $U(1)$ gauge field¹ via the interaction (1.3). The action is

$$\mathcal{S} = \int d^4x \sqrt{-g} \left[\frac{M_p^2}{2} R - \frac{1}{2} (\partial\varphi)^2 - V(\varphi) - \frac{1}{4} F^{\mu\nu} F_{\mu\nu} - \frac{\alpha}{4f} \varphi \tilde{F}^{\mu\nu} F_{\mu\nu} \right] \quad (2.1)$$

where R is the Ricci scalar, $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ the field strength, and $\tilde{F}^{\mu\nu} \equiv \frac{1}{2} \frac{\eta^{\mu\nu\alpha\beta}}{\sqrt{-g}} F_{\alpha\beta}$ its dual, with $\eta^{0123} = 1$. We separate the inflaton into a homogeneous (background) part plus its fluctuations

$$\varphi = \phi(t) + \delta\varphi(t, \vec{x}) \quad (2.2)$$

¹The generalization to non-Abelian gauge groups is straightforward.

We leave the potential $V(\varphi)$ arbitrary, except to assume that it is sufficiently flat to support the required amount of inflation ($N_e \gtrsim 60$). We assume a spatially flat Friedmann-Robertson-Walker (FRW) space-time with metric

$$ds^2 \equiv g_{\mu\nu} dx^\mu dx^\nu = -dt^2 + a^2(t) d\mathbf{x} \cdot d\mathbf{x} \quad (2.3)$$

$$= a^2(\tau) [-d\tau^2 + d\mathbf{x} \cdot d\mathbf{x}] \quad (2.4)$$

where on the second line we have introduced conformal time, τ , related to cosmic time as $ad\tau = dt$. Derivatives with respect to cosmic time are denoted as $\partial_t f \equiv \dot{f}$ and with respect to conformal time as $\partial_\tau f \equiv f'$. The Hubble rate $H \equiv \dot{a}/a$ has conformal time analogue $\mathcal{H} \equiv a'/a$.

We are first interested in the gauge quanta which are produced by the homogeneous rolling inflaton $\phi(\tau)$. To this end, we can ignore the inflaton and metric perturbations in the equations of motion of the gauge field (see section 5 for the complete treatment). Extremizing the action with respect to A_0 , and choosing the Coulomb gauge $A_0 = 0$, then gives $(\vec{\nabla} \cdot \vec{A})' = 0$, from which we set $\vec{\nabla} \cdot \vec{A} = 0$. The equations of motion for \vec{A} then read

$$\vec{A}'' - \nabla^2 \vec{A} - \frac{\alpha}{f} \phi' \vec{\nabla} \times \vec{A} = 0 \quad (2.5)$$

As we discuss in subsection 2.1, this equation describes the production of the quanta of the gauge fields that results from the motion of the inflaton.

The produced gauge quanta have two key effects: they backreact on the homogeneous background dynamics (see subsection 2.2) and also source inflaton perturbations (see subsection 2.3). Both effects are governed by the equation of motion of the inflaton, and the 00 Einstein equation, which read, respectively

$$\begin{aligned} \varphi'' + 2\mathcal{H}\varphi' - \nabla^2 \varphi + a^2 \frac{dV}{d\varphi} &= a^2 \frac{\alpha}{f} \vec{E} \cdot \vec{B} \\ \mathcal{H}^2 &= \frac{1}{3M_p^2} \left[\frac{1}{2} \varphi'^2 + \frac{1}{2} (\vec{\nabla} \varphi)^2 + a^2 V + \frac{a^2}{2} (\vec{E}^2 + \vec{B}^2) \right] \end{aligned} \quad (2.6)$$

In these equations we have retained the spatial dependence of φ (due to the inflaton perturbations), and we have introduced the physical “electric” and “magnetic” fields²

$$\vec{B} = \frac{1}{a^2} \vec{\nabla} \times \vec{A}, \quad \vec{E} = -\frac{1}{a^2} \vec{A}' \quad (2.7)$$

2.1 Production of Gauge Field Fluctuations

During inflation, the motion of the inflaton leads to an instability for the fluctuations of the gauge field. To see this effect, we start from the equation of motion for A_μ in the background of the homogeneous inflaton $\phi(t)$, eq. (2.5) above. We decompose the q-number field $\vec{A}(\tau, \mathbf{x})$ as

$$\vec{A}(\tau, \mathbf{x}) = \sum_{\lambda=\pm} \int \frac{d^3 k}{(2\pi)^{3/2}} \left[\vec{\epsilon}_\lambda(\mathbf{k}) a_\lambda(\mathbf{k}) A_\lambda(\tau, \mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{x}} + \text{h.c.} \right] \quad (2.8)$$

²We do not assume that A_μ necessarily corresponds to the Standard Model electro-magnetic gauge potential.

where “h.c.” denotes the Hermitian conjugate of the preceding term, the annihilation/creation operators obey

$$\left[a_\lambda(\mathbf{k}), a_{\lambda'}^\dagger(\mathbf{k}') \right] = \delta_{\lambda\lambda'} \delta^{(3)}(\mathbf{k} - \mathbf{k}') \quad (2.9)$$

Here $\vec{\epsilon}_\lambda$ are circular polarization vectors satisfying $\vec{k} \cdot \vec{\epsilon}_\pm(\vec{k}) = 0$, $\vec{k} \times \vec{\epsilon}_\pm(\vec{k}) = \mp i k \vec{\epsilon}_\pm(\vec{k})$, $\vec{\epsilon}_\pm(-\vec{k}) = \vec{\epsilon}_\pm(\vec{k})^*$, and normalized according to $\vec{\epsilon}_\lambda(\vec{k})^* \cdot \vec{\epsilon}_{\lambda'}(\vec{k}) = \delta_{\lambda\lambda'}$.

Inserting the decomposition (2.8) into eq. (2.5) results in the equation of motion

$$\left[\frac{\partial^2}{\partial \tau^2} + k^2 \pm \frac{2k\xi}{\tau} \right] A_\pm(\tau, k) = 0, \quad \xi \equiv \frac{\alpha \dot{\phi}}{2fH} \quad (2.10)$$

for the c-number mode functions A_\pm . During inflation the parameter ξ may be treated as constant, as its time variation is subleading in a slow roll expansion.

From equation (2.10) we see that one of the polarizations of \vec{A}_λ experiences a tachyonic instability for $k/(aH) \lesssim 2\xi$. Without loss of generality, we assume that $\dot{\phi} > 0$ during inflation, so that the mode exhibiting the instability is A_+ . In appendix A we review the solutions of (2.10) and show that the growth of fluctuations is well described by [15]

$$A_+(\tau, k) \cong \frac{1}{\sqrt{2k}} \left(\frac{k}{2\xi aH} \right)^{1/4} e^{\pi\xi - 2\sqrt{2\xi k/(aH)}} \quad (2.11)$$

in the interval $(8\xi)^{-1} \lesssim k/(aH) \lesssim 2\xi$ [18] of phase space that accounts for most of the power in the produced gauge fluctuations. The phase space of growing modes is non-vanishing for $\xi \gtrsim \mathcal{O}(1)$, which we assume throughout. Notice the exponential enhancement $e^{\pi\xi}$ in the solution (2.11), which arises due to tachyonic instability, and reflects significant nonperturbative gauge particle production in the regime $\xi \gtrsim 1$. On the other hand, the production of gauge field fluctuations is uninterestingly small for $\xi < 1$. Note also that the other polarization state, $A_-(\tau, k)$, is not produced and can therefore be ignored.

We have thus seen that the motion of the homogeneous inflaton $\phi(t)$ leads to production of gauge field quanta δA_μ . There are two key physical effects associated with the interactions of these produced quanta with the inflaton. The first effect is the backreaction of the produced quanta on the homogeneous dynamics of $\phi(t)$, $a(t)$. In the next subsection we study the conditions under which backreaction effects are negligible. The second key physical effect is the production of inflaton fluctuations via *inverse decay*; this is the subject of subsection 2.3.

2.2 Backreaction Effects

Backreaction effects can be accounted for using the mean of the field equations (2.6):

$$\ddot{\phi} + 3H\dot{\phi} + V'(\phi) = \frac{\alpha}{f} \langle \vec{E} \cdot \vec{B} \rangle \quad (2.12)$$

$$3H^2 = \frac{1}{M_p^2} \left[\frac{1}{2} \dot{\phi}^2 + V(\phi) + \frac{1}{2} \langle \vec{E}^2 + \vec{B}^2 \rangle \right] \quad (2.13)$$

where we have switched to physical time. The expectation values appearing in (2.12, 2.13) encode the backreaction of the produced gauge quanta on the homogeneous dynamics of

$\phi(t)$, $a(t)$. From (2.7) and (2.8), we have

$$\begin{aligned}\langle \vec{E} \cdot \vec{B} \rangle &= -\frac{1}{4\pi^2 a^4} \int dk k^3 \frac{d}{d\tau} |A_+|^2 \\ \frac{1}{2} \langle \vec{E}^2 + \vec{B}^2 \rangle &= \frac{1}{4\pi^2 a^4} \int dk k^2 [|A'_+|^2 + k^2 |A_+|^2]\end{aligned}\quad (2.14)$$

Since we are studying the backreaction of the produced quanta, we should disregard the modes that do not experience the growth discussed in the previous subsection, namely all modes A_- and the large momentum ($k > 2\xi aH$) modes of A_+ . As can be seen from (2.10), modes of A_+ with $k/(aH) \gg 2\xi$ remain in their vacuum state and do not experience any tachyonic instability. Such modes contribute to the vacuum energy of the $U(1)$ field (in eqn. (2.13); an analogous consideration applies to $\langle \vec{E} \cdot \vec{B} \rangle$ in eqn. (2.12)) that we assume is canceled by a bare vacuum energy, as is customary in QFT. (In short, we have nothing to add to the cosmological constant problem.) This prescription provides a UV cutoff $k/(aH) < 2\xi$ in the integrals (2.14); see the paragraph after eq. (2.10). From a direct inspection of the solutions of (2.10), we verified that the integrals (2.14) converge in the infrared $k \rightarrow 0$ region, and that they receive their support almost entirely from the region $(8\xi)^{-1} \lesssim k/(aH) \lesssim 2\xi$ in which (2.11) is a very good approximation of the exact solution. Therefore, (2.14) can be evaluated by using the expression (2.11) for A_+ , and by integrating only over this region of momenta. In fact, we verified that the momentum interval can be extended from 0 to ∞ , since the expression (2.11) rapidly decreases outside the $(8\xi)^{-1} \lesssim k/(aH) \lesssim 2\xi$ interval, and the contribution of the “outer” regions to (2.14) can be neglected. Proceeding in this way allows for an analytic result:

$$\langle \vec{E} \cdot \vec{B} \rangle \simeq -2.4 \cdot 10^{-4} \frac{H^4}{\xi^4} e^{2\pi\xi} \quad , \quad \left\langle \frac{\vec{E}^2 + \vec{B}^2}{2} \right\rangle \simeq 1.4 \cdot 10^{-4} \frac{H^4}{\xi^3} e^{2\pi\xi} \quad (2.15)$$

This procedure follows Ref. [15], where these expressions were first derived. We expect that any sensible renormalization prescription will yield results in agreement with (2.15).

From (2.12,2.13) we can distinguish two distinct kinds of backreaction effects. First, the gauge field fluctuations are produced at the expense of the kinetic energy of $\phi(t)$. This contributes a new source of dissipation into the homogeneous Klein-Gordon equation (2.12). In order to trust the usual slow-roll inflationary solution, we require that $|3H\dot{\phi}| \cong |-V'(\phi)| \gg |\frac{\alpha}{f} \langle \vec{E} \cdot \vec{B} \rangle|$. Using (2.15), this condition reads

$$\frac{H^2}{2\pi|\dot{\phi}|} \ll 13 \xi^{3/2} e^{-\pi\xi} \quad (2.16)$$

The second kind of backreaction effect arises because the energy density in produced gauge field fluctuations contributes to the Friedmann equation (2.13). To ensure that the expansion of the universe is dominated by the potential energy of the inflaton we require $3M_p^2 H^2 \cong V \gg \frac{1}{2} \langle \vec{E}^2 + \vec{B}^2 \rangle$. Once eq. (2.15) is taken into account, this condition reads

$$\frac{H}{M_p} \ll 146 \xi^{3/2} e^{-\pi\xi} \quad (2.17)$$

Taken together, the constraints (2.16,2.17) ensure that the unstable growth of gauge field fluctuations does not modify the usual homogeneous inflationary dynamics.

2.3 Inverse Decay Effects and Inflaton Perturbations

Even when (2.16, 2.17) are satisfied, the coupling $\varphi F \tilde{F}$ may still have a profound impact on the cosmological fluctuations in the model (2.1). The perturbations of the inflaton are described by the equation

$$\left[\frac{\partial^2}{\partial \tau^2} + 2\mathcal{H} \frac{\partial}{\partial \tau} - \nabla^2 + a^2 m^2 \right] \delta\varphi(\tau, \mathbf{x}) = a^2 \frac{\alpha}{f} \left(\vec{E} \cdot \vec{B} - \langle \vec{E} \cdot \vec{B} \rangle \right) \quad (2.18)$$

where $m^2 \equiv V''$. The solution of (2.18) splits into two parts: the solution of the homogeneous equation and the particular solution which is due to the source. Schematically, we have

$$\delta\varphi(\tau, \mathbf{x}) = \underbrace{\delta\varphi_{\text{vac}}(\tau, \mathbf{x})}_{\text{homogeneous}} + \underbrace{\delta\varphi_{\text{inv.decay}}(\tau, \mathbf{x})}_{\text{particular}} \quad (2.19)$$

The homogeneous solution corresponds, physically, to the usual vacuum fluctuations from inflation. The particular solution, on the other hand, can be interpreted as arising due to inverse decay processes $\delta A + \delta A \rightarrow \delta\varphi$. This new source of inflaton fluctuations contributes directly to the observable curvature perturbation on uniform density hypersurfaces, owing to the relation $\zeta \sim -\frac{H}{\phi} \delta\varphi$. It was shown in [18] that the inverse decay contribution to the cosmological fluctuations may actually *dominate* over the usual vacuum fluctuations in the regime $f \lesssim 10^{-2} M_p$. This radically modifies the phenomenology of axion inflation. In particular, the inverse decay contribution to the primordial cosmological fluctuations is highly nongaussian; this is evident already from (2.18) since the particular solution $\delta\varphi_{\text{inv.decay}}$ is bilinear in the gaussian field δA_μ .

To proceed with the computation, we decompose

$$\delta\varphi(\tau, \mathbf{x}) = \int \frac{d^3 k}{(2\pi)^{3/2}} \frac{Q_{\mathbf{k}}(\tau)}{a(\tau)} e^{i\mathbf{k} \cdot \mathbf{x}} \quad (2.20)$$

Eq. (2.18) then results in (we note that the last term of (2.18) has no effect on the mode functions with momentum different from zero) ³

$$\left[\partial_\tau^2 + k^2 + a^2 m^2 - \frac{a''}{a} \right] Q_{\mathbf{k}}(\tau) = J_{\mathbf{k}}(\tau) \quad (2.21)$$

$$J_{\mathbf{k}}(\tau) \equiv a^3(\tau) \frac{\alpha}{f} \int \frac{d^3 k}{(2\pi)^{3/2}} e^{-i\mathbf{k} \cdot \mathbf{x}} \vec{E} \cdot \vec{B} \quad (2.22)$$

We separate the mode functions of the two terms in eq. (2.19) as

$$Q_{\mathbf{k}}(\tau) = Q_{\mathbf{k}}^{\text{vac}}(\tau) + Q_{\mathbf{k}}^{\text{inv.decay}}(\tau) \quad (2.23)$$

and we discuss each contribution separately.

The homogeneous term is expanded as

$$Q_{\mathbf{k}}^{\text{vac}}(\tau) = b(\mathbf{k}) \varphi_k(\tau) + b^\dagger(-\mathbf{k}) \varphi_k^*(\tau) \quad (2.24)$$

³In Section 5 we present the complete computation, including also scalar metric perturbations. We show explicitly that equation (2.21) still holds, with only the addition of subdominant (Planck-suppressed) terms in the source $J_{\mathbf{k}}$.

The inflaton ladder operators obey

$$[b(\mathbf{k}), b^\dagger(\mathbf{k}')] = \delta^{(3)}(\mathbf{k} - \mathbf{k}') \quad (2.25)$$

and commute with the ladder operators of the gauge field

$$[b(\mathbf{k}), a_\lambda(\mathbf{k}')] = [b(\mathbf{k}), a_\lambda^\dagger(\mathbf{k}')] = 0 \quad (2.26)$$

The properly normalized homogeneous solutions of (2.21) are given by the well-known result

$$\varphi_k(\tau) = i \frac{\sqrt{\pi}}{2} \sqrt{-\tau} H_\nu^{(1)}(-k\tau), \quad \nu \cong \frac{3}{2} + \mathcal{O}(\epsilon, \eta) \quad (2.27)$$

where we have chosen the (arbitrary) phase so that $\varphi_k(\tau)$ is real in the limit $-k\tau \rightarrow 0$.

The vacuum modes (2.27) are employed in the retarded Green function associated with (2.21),

$$G_k(\tau, \tau') = i\Theta(\tau - \tau') [\varphi_k(\tau)\varphi_k^*(\tau') - \varphi_k^*(\tau)\varphi_k(\tau')] \quad (2.28)$$

which obeys $[\partial_\tau^2 + k^2 + a^2 m^2 - \frac{a''}{a}] G_k(\tau, \tau') = \delta(\tau - \tau')$.

Using the Green function (2.28) the particular solution of (2.21) takes the form

$$Q_{\mathbf{k}}^{\text{inv.decay}}(\tau) = \int_{-\infty}^0 d\tau' G_k(\tau, \tau') J_{\mathbf{k}}(\tau') \quad (2.29)$$

where the source term was defined in eq. (2.22). We note that this particular solution is statistically independent of the homogeneous solution (2.24). In fact, the particular solution can be expanded in terms of the annihilation/creation operators $a_\lambda(\mathbf{k}), a_\lambda^\dagger(\mathbf{k})$ associated with the gauge field, while the homogeneous solution can be expanded in terms of the annihilation/creation operators $b(\mathbf{k}), b^\dagger(\mathbf{k})$ associated with the inflaton vacuum fluctuations. As we pointed out, these two sets of operators commute with one another.

We are now in a position to compute the correlation functions for the perturbations $\delta\varphi$. We present this computation in section 3. In section 4 we discuss the resulting phenomenology. Finally, in section 5 we show that these result are valid also when metric perturbations are also consistently taken into account.

3. Correlation Functions

In this section we compute the main phenomenological signatures of the model (2.1). Specifically: in subsection 3.1 we compute the two point correlation function of the density perturbation $\zeta = -\frac{H}{\phi} \delta\varphi$; in subsection 3.2 we present the corresponding power spectrum; in subsection 3.3 we compute the three point correlation function of ζ ; in subsection 3.4 we present the corresponding bispectrum; in subsection 3.5 we finally summarize the computation of the power spectrum of the gravity waves (GW) modes. In this section we neglect scalar metric perturbations for simplicity, however, in section 5 we show explicitly that their consistent inclusion does not modify our results.

3.1 Two-Point Correlation Function

We start from the relation $\zeta(\tau, \vec{x}) = -\frac{H}{\dot{\phi}} \delta\varphi(\tau, \vec{x})$ between the inflaton perturbations, and the curvature perturbation on uniform density hypersurfaces.⁴ We decompose the latter as

$$\zeta(\tau, \mathbf{x}) = \int \frac{d^3k}{(2\pi)^{3/2}} \zeta_{\mathbf{k}}(\tau) e^{i\mathbf{k}\cdot\mathbf{x}} \quad (3.1)$$

so that $\zeta_{\mathbf{k}} = -\frac{H}{\dot{\phi}} \frac{Q_{\mathbf{k}}}{a}$. As we discussed in subsection 2.3, the inflaton perturbations comprise of two terms, one being the vacuum fluctuations, and one the fluctuations sourced by the gauge quanta; since these two terms are statistically independent, we have

$$\langle \zeta_{\mathbf{k}} \zeta_{\mathbf{k}'} \rangle = \frac{H^2}{a^2 \dot{\phi}^2} \left[\langle Q_{\mathbf{k}}^{\text{vac}} Q_{\mathbf{k}'}^{\text{vac}} \rangle + \langle Q_{\mathbf{k}}^{\text{inv.decay}} Q_{\mathbf{k}'}^{\text{inv.decay}} \rangle \right] \quad (3.2)$$

The contribution from the vacuum modes is standard. Using eqs. (2.24), (2.25), and (2.27), one obtains the well-known result

$$\begin{aligned} \langle \zeta_{\mathbf{k}}^{\text{vac}} \zeta_{\mathbf{k}'}^{\text{vac}} \rangle &= \frac{H^4}{2\dot{\phi}^2} \left(\frac{k}{aH} \right)^{n_s-1} \frac{1}{k^3} \delta^{(3)}(\mathbf{k} + \mathbf{k}') \\ &= \frac{2\pi^2}{k^3} \mathcal{P} \left(\frac{k}{aH} \right)^{n_s-1} \delta^{(3)}(\mathbf{k} + \mathbf{k}') \end{aligned} \quad (3.3)$$

in the late time / large scales limit, $-k\tau \ll 1$. In the second line of (3.3) we have introduced

$$n_s = 1 + 3 - 2\nu = 1 + \mathcal{O}(\epsilon, \eta) \quad (3.4)$$

$$\mathcal{P}^{1/2} \equiv \frac{H^2}{2\pi|\dot{\phi}|} \quad (3.5)$$

We now compute the 2-point correlator of the particular solution (2.29):

$$\langle \zeta_{\mathbf{k}}^{\text{inv.decay}}(\tau) \zeta_{\mathbf{k}'}^{\text{inv.decay}}(\tau) \rangle = \frac{H^2}{\dot{\phi}^2} \int d\tau' d\tau'' \frac{G_{\mathbf{k}}(\tau, \tau')}{a(\tau)} \frac{G_{\mathbf{k}'}(\tau, \tau'')}{a(\tau)} \langle J_{\mathbf{k}}(\tau') J_{\mathbf{k}'}(\tau'') \rangle \quad (3.6)$$

(In the following, we temporarily omit the superscript ‘inv.decay’ for notational convenience.) In evaluating this expression, we can make some approximations for the mode functions $\varphi_k(\tau)$ and $\varphi_k(\tau')$ appearing in the Green function (2.28). Since we are interested in the power spectrum of modes well outside the horizon we can use the small argument limit $-k\tau \ll 1$ for the modes (2.27) entering in the Green function (2.28),

$$\varphi_k(\tau) \simeq \frac{a(\tau)H}{\sqrt{2}} \frac{1}{k^{3/2}} (-k\tau)^{\frac{n_s-1}{2}}, \quad \text{for } -k\tau \ll 1 \quad (3.7)$$

Notice that we disregard the small slow roll corrections to the amplitude, but we retain them in the momentum dependence (since this controls the departure of the spectrum from scale invariance). This gives

$$\langle \zeta_{\mathbf{k}}(\tau) \zeta_{\mathbf{k}'}(\tau) \rangle \cong \frac{2H^4}{\dot{\phi}^2} \frac{(-k\tau)^{n_s-1}}{k^3} \int_{-\infty}^{\tau} d\tau' d\tau'' \text{Im}[\varphi_{\mathbf{k}}(\tau')] \text{Im}[\varphi_{\mathbf{k}'}(\tau'')] \langle J_{\mathbf{k}}(\tau') J_{\mathbf{k}'}(\tau'') \rangle \quad (3.8)$$

⁴See section 5 for more details on this relation.

In this relation, we have already used the fact that the correlator is nonvanishing only for $|\mathbf{k}| = |\mathbf{k}'|$. We stress that, in (3.8) we do *not* employ the approximation (3.7) for the modes $\varphi_k(\tau')$ which appear under the integral, since τ' must be integrated over.

To evaluate (3.8), we require the source correlator $\langle J_{\mathbf{k}}(\tau') J_{\mathbf{k}}(\tau'') \rangle$. Explicit evaluation gives

$$\begin{aligned} \langle J_{\mathbf{k}}(\tau') J_{\mathbf{k}'}(\tau'') \rangle &= \frac{\alpha^2 \delta^{(3)}(\mathbf{k} + \mathbf{k}')}{8 f^2 a(\tau') a(\tau'')} \int \frac{d^3 q}{(2\pi)^3} \left[1 + \frac{|\mathbf{q}|^2 - \mathbf{q} \cdot \mathbf{k}}{|\mathbf{q}| |\mathbf{k} - \mathbf{q}|} \right]^2 \\ &\quad \times \mathcal{A}[\tau', |\mathbf{q}|, |\mathbf{q} - \mathbf{k}|] \mathcal{A}^*[\tau'', |\mathbf{q}|, |\mathbf{q} - \mathbf{k}|] \end{aligned} \quad (3.9)$$

where

$$\mathcal{A}[\tau', |\mathbf{q}|, |\mathbf{q} - \mathbf{k}|] \equiv |\mathbf{q}| A'_+(\tau', |\mathbf{q} - \mathbf{k}|) A_+(\tau', |\mathbf{q}|) + |\mathbf{q} - \mathbf{k}| A'_+(\tau', |\mathbf{q}|) A_+(\tau', |\mathbf{q} - \mathbf{k}|) \quad (3.10)$$

and where the relations

$$|\vec{\epsilon}_+(\mathbf{q}) \cdot \vec{\epsilon}_+(\mathbf{k} - \mathbf{q})|^2 = \frac{1}{4} \left[1 + \frac{|\mathbf{q}|^2 - \mathbf{q} \cdot \mathbf{k}}{|\mathbf{q}| |\mathbf{k} - \mathbf{q}|} \right]^2, \quad \vec{\epsilon}_\lambda(-\mathbf{k}) = \vec{\epsilon}_\lambda^*(\mathbf{k}) \quad (3.11)$$

have been used.

Putting all together, we have

$$\begin{aligned} \langle \zeta_{\mathbf{k}}(\tau) \zeta_{\mathbf{k}'}(\tau) \rangle &\cong \frac{\alpha^2 H^6}{4 f^2 \dot{\phi}^2} \frac{(-k\tau)^{n_s-1}}{k^3} \delta^{(3)}(\mathbf{k} + \mathbf{k}') \int \frac{d^3 q}{(2\pi)^3} \left[1 + \frac{|\mathbf{q}|^2 - \mathbf{q} \cdot \mathbf{k}}{|\mathbf{q}| |\mathbf{k} - \mathbf{q}|} \right]^2 \\ &\quad \times \left| \int_{-\infty}^{\tau} d\tau' (-\tau') \text{Im}[\varphi_k(\tau')] \mathcal{A}[\tau', |\mathbf{q}|, |\mathbf{q} - \mathbf{k}|] \right|^2 \end{aligned} \quad (3.12)$$

Following similar arguments to those made in subsection 2.2, we can use the approximation (A-3) for the gauge field mode functions (we discuss this step more in details at the end of this subsection). This leads to

$$\begin{aligned} \langle \zeta_{\mathbf{k}}(\tau) \zeta_{\mathbf{k}'}(\tau) \rangle &\cong \frac{\alpha^2 H^6 e^{4\pi\xi}}{2^8 \pi^3 f^2 \dot{\phi}^2} \frac{(-k\tau)^{n_s-1}}{k^3} \delta^{(3)}(\mathbf{k} + \mathbf{k}') \\ &\quad \times \int d^3 q_* \left[1 + \frac{|\mathbf{q}_*|^2 - \mathbf{q}_* \cdot \hat{k}}{|\mathbf{q}_*| |\hat{k} - \mathbf{q}_*|} \right]^2 |\mathbf{q}_*|^{1/2} |\mathbf{q}_* - \hat{k}|^{1/2} \left[|\mathbf{q}_*|^{1/2} + |\mathbf{q}_* - \hat{k}|^{1/2} \right]^2 \\ &\quad \times \mathcal{I}^2 \left[2\sqrt{2\xi} \left(\sqrt{|\mathbf{q}_*|} + \sqrt{|\mathbf{q}_* - \hat{k}|} \right) \right] \end{aligned} \quad (3.13)$$

where the integration variable $\mathbf{q}_* \equiv \mathbf{q}/|\mathbf{k}|$ is dimensionless, and where

$$\mathcal{I}[z] \equiv \sqrt{\frac{\pi}{2}} \int_{-k\tau}^{\infty} dx x^{3/2} \text{Re} \left[H_{\nu}^{(1)}(x) \right] e^{-z\sqrt{x}} \quad (3.14)$$

(notice that $x \equiv -k\tau$). As we are interested only in super horizon modes, $-k\tau \ll 1$, we can set to zero the lower extreme of integration of \mathcal{I} . It is then manifest that we can set $\nu = 3/2$ in the argument of the Hankel function, since the slow roll corrections appearing

there only modify (in a negligible amount) the amplitude of the correlator, but not its scale dependence. This leads to

$$\mathcal{I}(z) \simeq \int_0^\infty dx (\sin x - x \cos x) e^{-z\sqrt{x}} \quad (3.15)$$

For future convenience, we rewrite the correlator (3.13) as

$$\langle \zeta_{\mathbf{k}}^{\text{inv.decay}}(\tau) \zeta_{\mathbf{k}'}^{\text{inv.decay}}(\tau) \rangle \equiv \frac{2\pi^2}{k^3} (-k\tau)^{n_s-1} \mathcal{P}^2 f_2(\xi) e^{4\pi\xi} \delta^{(3)}(\mathbf{k} + \mathbf{k}') \quad (3.16)$$

where \mathcal{P} was defined in (3.5) and the dimensionless function $f_2(\xi)$ is

$$\begin{aligned} f_2(\xi) \equiv & \frac{\xi^2}{8\pi} \int d^3q_* \left[1 + \frac{|\mathbf{q}_*|^2 - \mathbf{q}_* \cdot \hat{k}}{|\mathbf{q}_*| |\hat{k} - \mathbf{q}_*|} \right]^2 |\mathbf{q}_*|^{1/2} |\mathbf{q}_* - \hat{k}|^{1/2} \left[|\mathbf{q}_*|^{1/2} + |\mathbf{q}_* - \hat{k}|^{1/2} \right]^2 \\ & \times \mathcal{I}^2 \left[2\sqrt{2\xi} \left(\sqrt{|\mathbf{q}_*|} + \sqrt{|\mathbf{q}_* - \hat{k}|} \right) \right] \end{aligned} \quad (3.17)$$

In general, the function $f_2(\xi)$ needs to be evaluated numerically. However, a simplification is achieved when the argument of \mathcal{I} is much greater than one.

$$\mathcal{I}[z] \cong \int_0^\infty dx \frac{x^3}{3} e^{-z\sqrt{x}} = \frac{3360}{z^8} \quad \text{for } z \gg 1 \quad (3.18)$$

As $\sqrt{|\mathbf{q}_*|} + \sqrt{|\mathbf{q}_* - \hat{k}|} \geq 1$, this approximation is certainly appropriate at large ξ . One is then left with a two dimensional integral that numerically evaluates to

$$f_2(\xi) \cong \frac{7.5 \cdot 10^{-5}}{\xi^6}, \quad \xi \gg 1 \quad (3.19)$$

The degree of accuracy of this approximation can be seen in Figure 1. It is also useful to have a fit for f_2 in the range $2 \leq \xi \leq 3$, as this is the most relevant one for phenomenology (as we discussed in section 4). The best monomial fit to f_2 in this range is

$$f_2(\xi) \cong \frac{3 \cdot 10^{-5}}{\xi^{5.4}}, \quad 2 \leq \xi \leq 3 \quad (3.20)$$

We conclude this subsection with the justification of the use of (2.11) into the integral (3.12). We know that the expressions (2.11) are accurate approximations to the modes $A_+(\tau, k)$ in the range $1/8\xi \leq -\tau k \leq 2\xi$. As discussed in subsection 2.2, for any given mode k , this interval corresponds to the times for which the amplification of the gauge field is maximal. The key point is to ensure that there is a common region in the integration space of (3.12) for which both the functions A_+ that enter in the expression of \mathcal{A} can be approximated by (2.11). This requires that

$$\frac{1}{8\xi} \leq -|\mathbf{q}| \tau' \leq 2\xi \quad \text{and} \quad \frac{1}{8\xi} \leq -|\mathbf{q} - \mathbf{k}| \tau' \leq 2\xi \quad (3.21)$$

or, equivalently,

$$\frac{1}{8\xi} \leq |\mathbf{q}_*| x \leq 2\xi \quad \text{and} \quad \frac{1}{8\xi} \leq |\mathbf{q}_* - \hat{k}| x \leq 2\xi \quad (3.22)$$

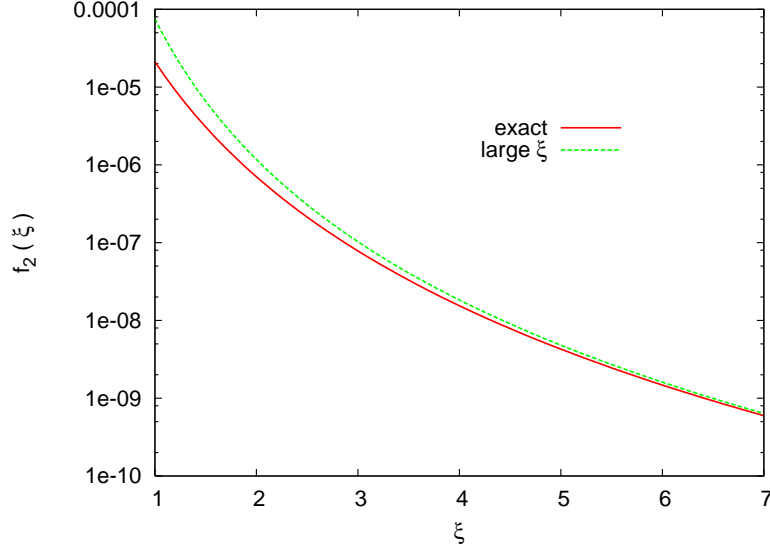


Figure 1: Exact evaluation and large argument approximation of the function f_2 .

($x = -k\tau'$ is the integration variable in (3.14), and $\mathbf{q}_* = \mathbf{q}/|\mathbf{k}|$). For the super-horizon modes that are relevant for phenomenology, we know that x extends from 0 to ∞ . Therefore, for any value of \mathbf{q}_* , there are always values of the rescaled time x for which either $A_+(\tau', |\mathbf{q}_*|)$ or $A_+(\tau', |\mathbf{q}_* - \hat{k}|)$ is maximal, and approximated by (A-3). Only when $|\mathbf{q}_*| \simeq |\mathbf{q}_* - \hat{k}|$, both conditions are valid for the same values of x , and the approximations (2.11) can be used in the whole integrand of (2.2). Therefore, our result is correct only if the integrand of (2.2) is strongly peaked at $|\mathbf{q}_*| = \mathcal{O}(1)$. We have verified with direct inspection that this is indeed the case (We have verified this claim also using the representation (A-2) of the gauge field modes which is valid arbitrarily deep in the IR; see Appendix C.)

There is a clear physical reason why the integrand is strongly peaked in the region of phase space where the conditions (3.21) are satisfied. Notice that, for the values $z = \mathcal{O}(1)$ which are relevant for the present computation, the expression (3.15) has most of its support at $x = \mathcal{O}(1)$. This means that the “imprint” of the fluctuations ζ_k from the gauge modes occurs when the wavelength of the fluctuation is of the order of the horizon scale (x of order one), and it is caused by the inverse decay of gauge field modes, whose wavelength is also of the order of the horizon scale ($|\mathbf{q}_*|$ of order one). This is a natural outcome of causality/locality. Identical considerations apply in the computation of the three point function that we perform in subsection 3.3.

3.2 The Primordial Power Spectrum

The two point correlation function in momentum space is related to the power spectrum by the standard expression

$$\langle \zeta_{\mathbf{k}} \zeta_{\mathbf{k}'} \rangle \equiv P_\zeta(k) \frac{2\pi^2}{k^3} \delta^{(3)}(\mathbf{k} + \mathbf{k}') \quad (3.23)$$

As we have seen in the previous subsection, $\zeta_{\mathbf{k}} = \zeta_{\mathbf{k}}^{\text{vac}} + \zeta_{\mathbf{k}}^{\text{inv.decay}}$, and the two terms are uncorrelated. From (3.3) and (3.16) we get the power spectrum at late times

$$P_{\zeta}(k) = \mathcal{P} \left(\frac{k}{k_0} \right)^{n_s-1} \left[1 + \mathcal{P} f_2(\xi) e^{4\pi\xi} \right] \quad (3.24)$$

where $\mathcal{P}^{1/2} = \frac{H^2}{2\pi|\phi|}$.

3.3 Three-Point Correlation Function

In this subsection we calculate the three point function of $\zeta_{\mathbf{k}}^{\text{inv.decay}}$,

$$\begin{aligned} \langle \zeta_{\mathbf{k}_1}(\tau) \zeta_{\mathbf{k}_2}(\tau) \zeta_{\mathbf{k}_3}(\tau) \rangle &= -\frac{H^3}{\dot{\phi}^3} \int d\tau_1 d\tau_2 d\tau_3 \frac{G_{k_1}(\tau, \tau_1)}{a(\tau)} \frac{G_{k_2}(\tau, \tau_2)}{a(\tau)} \frac{G_{k_3}(\tau, \tau_3)}{a(\tau)} \\ &\quad \times \langle J_{\mathbf{k}_1}(\tau_1) J_{\mathbf{k}_2}(\tau_2) J_{\mathbf{k}_3}(\tau_3) \rangle \end{aligned} \quad (3.25)$$

Proceeding as in subsection 3.1, we arrive to

$$\begin{aligned} \langle \zeta_{\mathbf{k}_1} \zeta_{\mathbf{k}_2} \zeta_{\mathbf{k}_3} \rangle &= -\frac{\alpha^3 H^9}{f^3 \dot{\phi}^3 k_1^3 k_2^3 k_3^3} \delta^{(3)}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) \int \frac{d^3 q}{(2\pi)^{9/2}} \\ &\quad \times [\vec{\epsilon}(\mathbf{q}) \cdot \vec{\epsilon}(\mathbf{k}_1 - \mathbf{q})] [\vec{\epsilon}(\mathbf{q} - \mathbf{k}_1) \cdot \vec{\epsilon}(-\mathbf{q} - \mathbf{k}_3)] [\vec{\epsilon}(\mathbf{q} + \mathbf{k}_3) \cdot \vec{\epsilon}(-\mathbf{q})] \\ &\quad \times \Pi_i \int_{-\infty}^0 d\tau_i [k_i \tau_i \cos(k_i \tau_i) - \sin(k_i \tau_i)] \\ &\quad \times \mathcal{A}[\tau_1, |\mathbf{q}|, |\mathbf{k}_1 - \mathbf{q}|] \mathcal{A}[\tau_2, |\mathbf{k}_1 - \mathbf{q}|, |\mathbf{k}_3 + \mathbf{q}|] \mathcal{A}[\tau_3, |\mathbf{k}_3 + \mathbf{q}|, |\mathbf{q}|] \end{aligned} \quad (3.26)$$

where \mathcal{A} was defined in (3.10). In this expression we have used the fact that the mode functions A_+ are real (which is true both in the approximations (A-2) and (2.11)), and we have disregarded any (mild) scale dependence.⁵

The correlator depends on the size and the shape of the triangle formed by the vectors \mathbf{k}_i . We denote

$$|\mathbf{k}_1| = k, \quad |\mathbf{k}_2| = x_2 k, \quad |\mathbf{k}_3| = x_3 k \quad (3.27)$$

and, for future convenience, we parametrize the correlator as

$$\begin{aligned} \langle \zeta_{\mathbf{k}_1} \zeta_{\mathbf{k}_2} \zeta_{\mathbf{k}_3} \rangle &\equiv \frac{3}{80 (2\pi)^{7/2}} \frac{\alpha^3 H^9}{f^3 \dot{\phi}^3} \frac{e^{6\pi\xi}}{\xi^3} \frac{\delta^{(3)}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3)}{k^6} \frac{1 + x_2^3 + x_3^3}{x_2^3 x_3^3} f_3(\xi, x_2, x_3) \\ &= \frac{3}{10} (2\pi)^{5/2} \mathcal{P}^3 e^{6\pi\xi} \frac{\delta^{(3)}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3)}{k^6} \frac{1 + x_2^3 + x_3^3}{x_2^3 x_3^3} f_3(\xi, x_2, x_3) \end{aligned} \quad (3.28)$$

We proceed from eq. (3.26) as we did in subsection 3.1. Using the last expression, eq. (3.10), and eqs. (A-3), we arrive to

$$\begin{aligned} f_3(\xi; x_2, x_3) &= \frac{5}{3\pi} \frac{\xi^3}{x_2 x_3 [1 + x_2^3 + x_3^3]} \int d^3 q_* |\mathbf{q}_*|^{1/2} |\hat{k}_1 - \mathbf{q}_*|^{1/2} |\mathbf{q}_* + x_3 \hat{k}_3|^{1/2} \\ &\quad \times \left[\vec{\epsilon}(\mathbf{q}_*) \cdot \vec{\epsilon}(\hat{k}_1 - \mathbf{q}_*) \right] \left[|\mathbf{q}_*|^{1/2} + |\hat{k}_1 - \mathbf{q}_*|^{1/2} \right] \mathcal{I} \left[2\sqrt{2\xi} \left(|\mathbf{q}_*|^{1/2} + |\hat{k}_1 - \mathbf{q}_*|^{1/2} \right) \right] \end{aligned}$$

⁵Specifically, we have set $\nu = 3/2$ in the mode functions entering in the Green function (2.28)).

$$\begin{aligned}
& \times \left[\vec{\epsilon}(\mathbf{q}_* - \hat{k}_1) \cdot \vec{\epsilon}(-\mathbf{q}_* - x_3 \hat{k}_3) \right] \left[|\hat{k}_1 - \mathbf{q}_*|^{1/2} + |\mathbf{q}_* + x_3 \hat{k}_3|^{1/2} \right] \\
& \quad \mathcal{I} \left[2\sqrt{\frac{2\xi}{x_2}} \left(|\hat{k}_1 - \mathbf{q}_*|^{1/2} + |\mathbf{q}_* + x_3 \hat{k}_3|^{1/2} \right) \right] \\
& \times \left[\vec{\epsilon}(\mathbf{q}_* + x_3 \hat{k}_3) \cdot \vec{\epsilon}(-\mathbf{q}_*) \right] \left[|\mathbf{q}_* + x_3 \hat{k}_3|^{1/2} + |\mathbf{q}_*|^{1/2} \right] \\
& \quad \mathcal{I} \left[2\sqrt{\frac{2\xi}{x_3}} \left(|\mathbf{q}_* + x_3 \hat{k}_3|^{1/2} + |\mathbf{q}_*|^{1/2} \right) \right] \quad (3.29)
\end{aligned}$$

To continue, we can set

$$\begin{aligned}
\hat{k}_1 &= (1, 0, 0) \\
x_3 \hat{k}_3 &= -\frac{1}{2} \left(1 - x_2^2 + x_3^2, \sqrt{-(1 - x_2 + x_3)(1 + x_2 - x_3)(1 - x_2 - x_3)(1 + x_2 + x_3)}, 0 \right)
\end{aligned} \quad (3.30)$$

and we note that, to a generic vector $\mathbf{k} = k(\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)$, corresponds the polarization operator

$$\vec{\epsilon}_+(\mathbf{k}) = \frac{1}{\sqrt{2}} (\cos \theta \cos \phi - i \sin \phi, \cos \theta \sin \phi + i \cos \phi, -\sin \theta) \quad (3.31)$$

(it is immediate to verify that this expression satisfies all the properties listed after eq. (2.9)).

In the reminder of this subsection we evaluate f_3 for the equilateral configuration $x_2 = x_3 = 1$. In the large ξ limit, we can use the analytic approximation (3.18) for \mathcal{I} , and perform the momentum integral numerically. We obtain

$$f_3(\xi; 1, 1) \cong \frac{2.8 \cdot 10^{-7}}{\xi^9} \quad , \quad \xi \gg 1 \quad (3.32)$$

The degree of accuracy of this approximation can be seen in Figure 2. It is also useful to have a fit for f_3 in the range $2 \leq \xi \leq 3$, as this is the most relevant one for phenomenology (as we discuss in section 4). The best monomial fit to f_3 in this range is

$$f_3(\xi; 1, 1) \cong \frac{7.4 \cdot 10^{-8}}{\xi^{8.1}} \quad , \quad 2 \leq \xi \leq 3 \quad (3.33)$$

3.4 The Bispectrum and Nonlinearity Parameter

A popular parametrization of nongaussianity is the nonlinearity parameter f_{NL} , introduced by assuming that the curvature perturbation may be expanded as

$$\zeta(\mathbf{x}) = \zeta_g(\mathbf{x}) + \frac{3}{5} f_{NL} [\zeta_g^2(\mathbf{x}) - \langle \zeta_g^2(\mathbf{x}) \rangle] \quad (3.34)$$

where $\zeta_g(x)$ is a gaussian random field (see [35] for a careful discussion of sign conventions). Both ζ and ζ_g may be decomposed as in (3.1) so that the relation between the q-modes of the Fourier decomposition is

$$\zeta_{\mathbf{k}} = \zeta_{g,\mathbf{k}} + \frac{3}{5} f_{NL} \int \frac{d^3 p}{(2\pi)^{3/2}} \zeta_{g,\mathbf{k}} \zeta_{g,\mathbf{k}-\mathbf{p}} \quad (3.35)$$

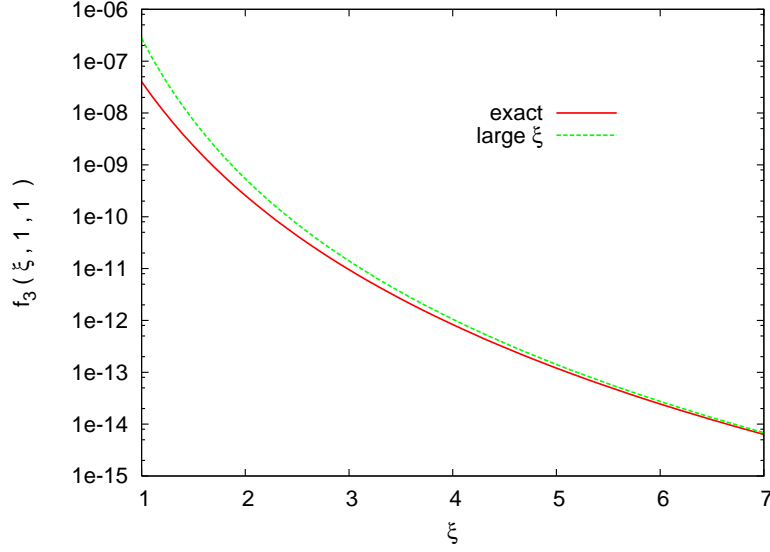


Figure 2: Exact evaluation and large argument approximation of the function $f_3(\xi; 1, 1)$.

By definition, the three point correlator of ζ_g vanishes. However, due to the quadratic term in (3.34), the three point correlator of ζ is nonvanishing, and can be expressed through a sum of two point correlators of ζ_g . One finds⁶

$$\langle \zeta_{\mathbf{k}_1} \zeta_{\mathbf{k}_2} \zeta_{\mathbf{k}_3} \rangle = \frac{3}{10} (2\pi)^{5/2} f_{\text{NL}} P_\zeta(k)^2 \delta^{(3)}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) \frac{\sum_i k_i^3}{\prod_i k_i^3} \quad (3.36)$$

where the power spectrum was defined in (3.23). To obtain this expression, recall that the ladder operators are normalized according to (2.25), and that the power spectrum $P(k)$ is related to the two point function as in (3.23). One should also identify the two point function of ζ with that of ζ_g (as the difference in subleading in a perturbative expansion), and disregard the mild scale dependence of the power spectrum.

By comparison to (3.36), one may define an “effective” (momentum dependent) non-linearity parameter, even when the intrinsic nongaussianity is not of the local form (3.34). For axion inflation, using the parametrization (3.28) of the three point correlator, we can write

$$f_{\text{NL}}^{\text{eff}}(\xi; x_2, x_3) = \frac{f_3(\xi; x_2, x_3) \mathcal{P}^3 e^{6\pi\xi}}{P_\zeta(k)^2} \quad (3.37)$$

where we recall that $\mathcal{P}^{1/2} = \frac{H^2}{2\pi|\phi|}$.

3.5 Power spectrum of the Tensor Modes

The produced gauge quanta also source tensor metric perturbations (gravity waves). The total GW power spectrum was first given in [18]. Ref. [36] then pointed out that the chiral nature of the GW produced by the gauge modes can be probed through the resulting

⁶Note that the factors of $(2\pi)^{n/2}$ differ from [23]. This stems from a different convention for the normalization of the Fourier transform (3.1).

nonvanishing $\langle BE \rangle$ and $\langle BT \rangle$ correlators of the CMB, as studied in [37, 38]. For the present model, a positive detection of parity violation would only be possible in a cosmic variance limited experiment, and for a limited portion of the parameter space [36]. In particular, one needs to be in a regime in which the GW production from the gauge modes dominates over that from the vacuum. In the minimal version of the model studied here, this region is ruled out by the nongaussianity limit that we discuss in the next section [18]. Ref. [36] circumvented this problem by considering the presence of many ($\mathcal{N} \gtrsim 10^3$) gauge fields, or by the use of a curvaton.

In the short paper [18] we only reported the final result, reserving the present work for the presentation of the details of the computation. As in the meantime this computation has been presented in details in ref. [36], we only provide a quick summary here. The tensor modes enter in the spatial components of the metric as $g_{ij} = a^2 (\delta_{ij} + h_{ij})$, with $h_{ii} = \partial_i h_{ij} = 0$. From the Einstein equations, one then finds

$$\frac{1}{2a^2} \left(\partial_\tau^2 + \frac{2a'}{a} \partial_\tau - \partial_{\mathbf{x}}^2 \right) h_{ij} = \frac{1}{M_p^2} (-E_i E_j - B_i B_j)^{TT} \quad (3.38)$$

where TT denotes the transverse and traceless projection of the spatial components of the energy-momentum tensor of the gauge field. The computation of the GW production is performed analogously to that of the density perturbations that we have presented in details in the previous subsections. As for density perturbations, the GW modes produced by the gauge quanta are uncorrelated with those from the vacuum, and the two contributions add up incoherently in the power spectrum. The two GW helicities are obtained from the projectors $\Pi_{ij,L/R}(\mathbf{k}) = \epsilon_i^{(\mp)}(\mathbf{k}) \epsilon_j^{(\mp)}(\mathbf{k})$. One finds the two power spectra

$$P_{L/R} \cong \frac{H^2}{\pi^2 M_p^2} \left(\frac{k}{k_0} \right)^{n_T} \left[1 + \frac{2H^2}{M_p^2} f_{h,L/R}(\xi) e^{4\pi\xi} \right] \quad (3.39)$$

where

$$f_{h,L/R} = \frac{1}{\xi} \int \frac{d^3 q_*}{(2\pi)^3} \sqrt{q_* |\hat{k} - \mathbf{q}_*|} \frac{(1 \pm \cos \theta)^2 \left(1 - q_* \cos \theta \pm \sqrt{1 - 2q_* \cos \theta + q_*^2} \right)^2}{16 (1 - 2q_* \cos \theta + q_*^2)} \times \\ \left\{ \int_0^\infty dx \sqrt{x} [\sin x - x \cos x] \left[\frac{2\xi}{x} + \sqrt{q_* |\hat{k} - \mathbf{q}_*|} \right] e^{-2\sqrt{2\xi}x} \left[\sqrt{\tilde{q}_*} + \sqrt{|\hat{k} - \tilde{q}_*|} \right] \right\}^2 \quad (3.40)$$

and where q_* and θ are, respectively, the magnitude of the (dimensionless) integration momentum \mathbf{q}_* , and the angle between this vector and the momentum \mathbf{k} of the mode.

In eq. (3.39), $n_T = -2\epsilon$. We note that the contribution to the spectrum of the modes from the vacuum, and of the modes sourced from the gauge field have the same scale dependence. The reason for this is that the scale dependence of the second term originates from the homogeneous solutions of (3.38) - which are the vacuum solutions - employed in the Green function. For the same reason, also the two terms in the scalar power spectrum have identical scale dependence, see subsection 3.1.

At large ξ , the integral over x can be performed as in the previous subsections, and one finds [36]

$$f_{h,L} \cong \frac{4.3 \cdot 10^{-7}}{\xi^6}, \quad f_{h,R} \cong \frac{9.2 \cdot 10^{-10}}{\xi^6}, \quad \xi \gg 1 \quad (3.41)$$

The degree of accuracy of this approximation can be seen in Figure 3. It is also useful to have a fit for $f_{h,L}$ in the range $2 \leq \xi \leq 3$, as this is the most relevant one for phenomenology (as we discussed in section 4). The best monomial fit to $f_{h,L}$ in this range is

$$f_{h,L}(\xi) \cong \frac{2.6 \cdot 10^{-7}}{\xi^{5.7}} \quad , \quad 2 \leq \xi \leq 3 \quad (3.42)$$

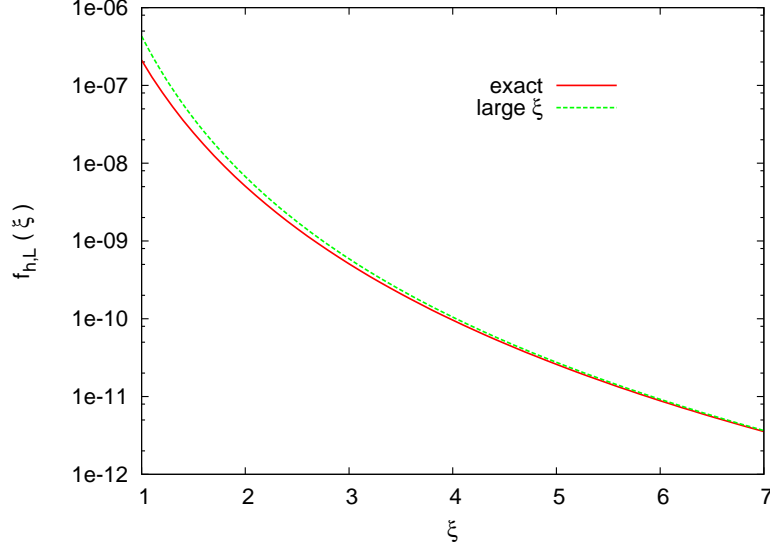


Figure 3: Exact evaluation and large argument approximation of the function $f_{h,L}(\xi)$.

From (3.41) we see that, for the purpose of computing the tensor-to-scalar ratio, one can disregard the right helicity GW modes produced by the gauge fields. One then obtains the result

$$P_{GW} = P_{h,L} + P_{h,R} \cong \frac{2H^2}{\pi^2 M_p^2} \left(\frac{k}{k_0} \right)^{n_T} \left[1 + \frac{H^2}{M_p^2} f_{h,L}(\xi) e^{4\pi\xi} \right] \quad (3.43)$$

first reported in [18].

4. Phenomenology of Axion Inflation

From the observational perspective, the key quantities which characterize any model of inflation are the spectrum of scalar and tensor perturbations, P_ζ and P_{GW} , along with the bispectrum of scalar perturbations B_ζ that encodes the leading departures from gaussian statistics. The explicit computations of these quantities was presented in Section 3. In this Section, we study the resulting observational signatures.

4.1 COBE Normalization and Spectral Tilt

In subsections 3.1 and 3.2 we found that two uncorrelated terms contribute to the power spectrum in axion inflation. These are the usual fluctuations generated from the vacuum, along with the modes produced by the inverse decay of the gauge quanta excited by the

motion of the inflaton. Taking into account both contributions results in a power spectrum of the form

$$P_{\zeta}(k) = \mathcal{P} \left(\frac{k}{k_0} \right)^{n_s-1} \left[1 + \mathcal{P} f_2(\xi) e^{4\pi\xi} \right] \quad (4.1)$$

where $n_s = 1 + 2\eta - 6\epsilon$ is the spectral index, the pivot is $k_0 = 0.002 \text{ Mpc}^{-1}$, and we have defined

$$\mathcal{P}^{1/2} \equiv \frac{H^2}{2\pi|\dot{\phi}|} \quad (4.2)$$

It is worth noting that both terms in (4.1) have the *same* scale dependence. Thus, we recover the standard prediction for the scalar spectral tilt in single field inflation. The function $f_2(\xi)$ which appears in (4.1) is plotted in Fig. 1. For $\xi \gg 1$ eq. (3.19) provides an asymptotic expression for f_2 while, on the other hand, eq. (3.20) provides a good fit in the $2 \leq \xi \leq 3$ interval (as we discuss below, this is the most interesting interval for phenomenology). The COBE normalization $P_{\zeta}(k) \cong 25 \cdot 10^{-10}$ is satisfied along the curve

$$\mathcal{P}_{\text{COBE}}(\xi) \cong \frac{e^{-4\pi\xi}}{2f_2(\xi)} \left[-1 + \sqrt{1 + 10^{-8} f_2(\xi) e^{4\pi\xi}} \right] \quad (4.3)$$

in the $\xi - \mathcal{P}$ plane. We see that, at low ξ , the contribution to (4.1) sourced by the gauge field is subdominant and we recover the standard result $\mathcal{P}_{\text{COBE}}^{1/2} \cong 5 \cdot 10^{-5}$. The two contributions in (4.1) become equal at $\xi \cong 2.9$. As ξ is increased, inverse decay effects dominate the spectrum and the value of $\mathcal{P}_{\text{COBE}}^{1/2}$ must be (exponentially) decreased to avoid over-producing density perturbations.

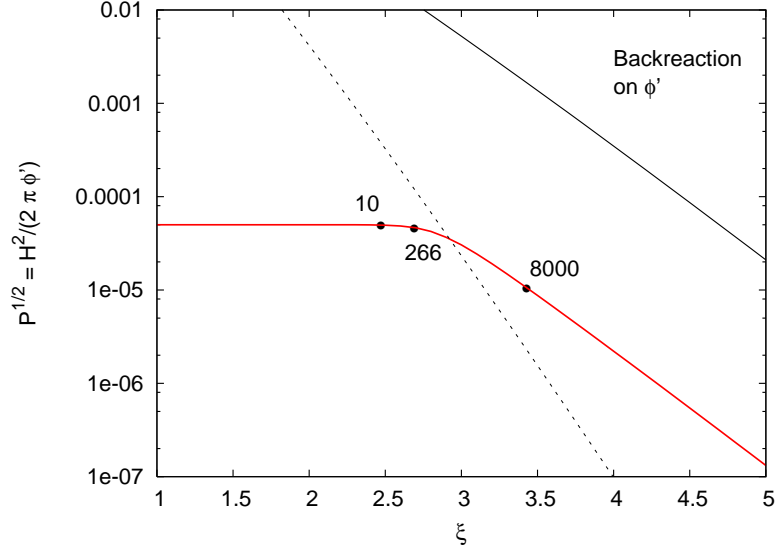


Figure 4: Values of parameters leading to the observed COBE normalization of the power spectrum (red line), and reference values for the nongaussianity parameter $f_{NL}^{\text{equil}} = 10, 266, 8000$ along this curve. See the main text for details.

The curve (4.3) is shown (red solid line) in Figure 4. The black dashed line shown in the Figure separates the (lower) region in which the vacuum fluctuations dominate from

the (upper) region in which the fluctuations from inverse decay dominate. In the region above the solid black line in Fig. 4 the backreaction bound (2.16) is violated. In that region of parameter space, the production of gauge field fluctuations is so strong that dissipative effects (rather than the potential $V(\phi)$) dominate the motion of $\dot{\phi}$. From Fig. 4 we see that this backreaction effect can be safely disregarded after the COBE normalization is imposed.⁷ Finally, Figure 4 shows some reference values of f_{NL}^{equil} along the COBE normalized curve; we discuss this in subsection 4.3.

4.2 Tensor-to-Scalar Ratio

Similarly to the result (4.1), the GW spectrum is also the sum of two uncorrelated contributions, one due the modes generated from the vacuum, and one due to the modes sources by the gauge field quanta. In subsection 3.5 we derived the result

$$P_{GW} \cong \frac{2H^2}{\pi^2 M_p^2} \left(\frac{k}{k_0} \right)^{n_T} \left[1 + \frac{H^2}{M_p^2} f_{h,L}(\xi) e^{4\pi\xi} \right] \quad (4.4)$$

where $n_T = -2\epsilon$. The second term in the square braces corresponds to the gravitational waves sourced by gauge field quanta. The function $f_{h,L}$ is plotted in Figure 3. Eq. (3.41) provides a large argument expansion while eq. (3.42) provides a fit in the $2 \leq \xi \leq 3$ interval (as we discuss below, this is the most interesting interval for phenomenology).

We define the tensor-to-scalar ratio in the usual way, by normalizing the amplitude of the power in gravitational waves to that in scalar fluctuations

$$r \equiv \frac{P_{GW}}{P_\zeta} \quad (4.5)$$

For $\xi \lesssim 1$, inverse decay effects are negligible and we recover the standard consistency relation for r , familiar from single field inflation. At $\xi \rightarrow \infty$, on the other hand, r tends to a different constant value, which is smaller than the asymptotic $\xi \rightarrow 0$ value. From (4.1), (4.4), and from the result $n_T = -2\epsilon$ it is easy to show that

$$r = \begin{cases} -8 n_T & \text{if } \xi \lesssim 3; \\ 1.8 n_T^2 & \text{as } \xi \rightarrow \infty. \end{cases} \quad (4.6)$$

Therefore, r is independent of ξ if either vacuum fluctuations or inverse decay effects dominate, while it interpolates between these two asymptotic values for intermediate ξ .

4.3 Nongaussianity

As discussed in section 2, the cosmological fluctuations generated by inverse decay effects are highly nongaussian. There are many different ways to parametrize departures from gaussianity. A standard work-horse is the local ansatz:

$$\zeta(x) = \zeta_g(x) + \frac{3}{5} f_{NL}^{\text{local}} [\zeta_g^2(x) - \langle \zeta_g^2(x) \rangle] \quad (4.7)$$

⁷One should also ensure that the energy density of the produced quanta gives a negligible contribution to the expansion of the universe. The resulting condition, eq. (2.17), cannot be shown in the $\xi - \mathcal{P}$ plane, and therefore needs to be studied on a case by case basis. We have verified that this condition is satisfied for the models studied in this section.

where ζ_g is a gaussian random field and f_{NL}^{local} quantifies the amount of nongaussianity. Although this simple parametrization has received considerable attention, it is certainly not the only well-motivated model for a nongaussian curvature perturbation. More generally, one should consider the bispectrum, $B_\zeta(k_i)$, which is the 3-point correlation function of ζ in Fourier space:

$$\langle \zeta_{\mathbf{k}_1} \zeta_{\mathbf{k}_2} \zeta_{\mathbf{k}_3} \rangle = B_\zeta(k_i) \delta^{(3)}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) \quad (4.8)$$

The bispectrum is a function of three momenta, \mathbf{k}_i , that form a triangle: $\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 = 0$. Hence a generic bispectrum may be characterized by specifying three interesting properties: the magnitude of the function, its dependence on the shape of the triangle, and its dependence of the size of the triangle. These properties are usually referred to as the *size*, *shape* and *running* of the nongaussianity, respectively. To characterize these properties, we find it convenient to introduce

$$|\mathbf{k}_1| = k \quad , \quad |\mathbf{k}_2| = x_2 k \quad , \quad |\mathbf{k}_3| = x_3 k \quad (4.9)$$

so that k encodes the overall size of the triangle while the dimensionless quantities x_2, x_3 encode its shape.

If we assume a local ansatz (4.7) then the bispectrum has a very particular dependence on momenta:

$$\begin{aligned} B_\zeta^{\text{local}}(k_i) &= \frac{3}{10} (2\pi)^{5/2} f_{NL}^{\text{local}} P_\zeta(k)^2 \frac{\sum_i k_i^3}{\prod_i k_i^3} \\ &= \frac{3(2\pi)^{5/2}}{10} \frac{P_\zeta(k)^2}{k^6} \frac{1 + x_2^3 + x_3^3}{x_2^3 x_3^3} f_{NL}^{\text{local}} \end{aligned} \quad (4.10)$$

This function peaks in the squeezed limit where one of the wave-numbers is much smaller than the other two (for example $k_1 \ll k_2, k_3$).

4.3.1 The Size and Running of the Nongaussianity

The bispectrum from axion inflation contains two uncorrelated contributions corresponding, respectively, to the usual vacuum fluctuations and to the fluctuations generated by inverse decay processes. As is well known, the former contribution gives rise to undetectably small nongaussianity and may be ignored. The second contribution, however, is more interesting and this was computed in subsection 3.4. We found that the bispectrum from axion inflation is very different from the local form (4.10). The bispectrum from axion inflation peaks on equilateral, rather than squeezed, triangles. Nevertheless, it is conventional to characterize the size of nongaussianity by matching to (4.10) on equilateral triangles $|\mathbf{k}_1| = |\mathbf{k}_2| = |\mathbf{k}_3|$. Proceeding in this way we find

$$f_{NL}^{\text{equil}} \equiv \frac{f_3(\xi; 1, 1) \mathcal{P}^3 e^{6\pi\xi}}{P_\zeta(k)^2} \quad (4.11)$$

The function $f_3(\xi; 1, 1)$ is plotted in Figure 2. Eq. (3.32) provides a large argument expansion of this function, while eq. (3.33) provides a fit in the $2 \leq \xi \leq 3$ interval.

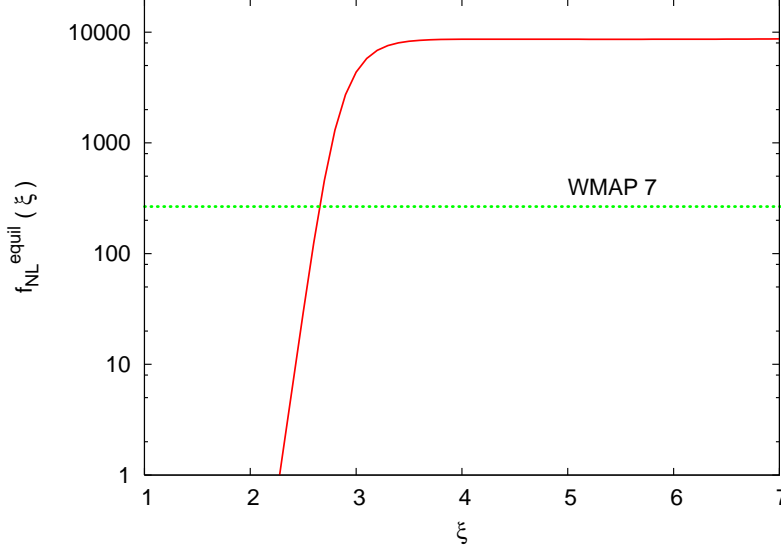


Figure 5: Value of $f_{NL}^{\text{equil}}(\xi)$ due to the perturbations ζ sourced by the gauge quanta. Also shown is the 95% CL upper bound from WMAP 7 [39].

We stress that this result does not include the negligible contribution from $\langle \delta\phi_{\text{vac}}^3 \rangle$ and is accurate as long as $|f_{NL}^{\text{equil}}| \gtrsim 1$.

From eqs. (4.3) and (4.11), from the observed power $P_\zeta \cong 25 \cdot 10^{-10}$ (we disregard the slow-roll suppressed scale dependence of f_{NL}), and from the expressions of f_2 and f_3 computed in the previous section, we can compute the value of $f_{NL}^{\text{equil}}(\xi)$ along the COBE normalized curve. We show this in Figure 5 (see also the reference values shown in Figure 4). We notice that $f_{NL}^{\text{equil}}(\xi)$ saturates to a constant value at large ξ (in the region where the vacuum contribution to P_ζ is negligible). From the large value asymptotic expressions (3.19) and (3.32), we find $f_{NL}^{\text{equil}}(\infty) \cong 8,600$. This value is already above the 95% CL upper $-214 < f_{NL}^{\text{equil}} < 266$ obtained from the WMAP 7 data [39]. This limit rules out $\xi \gtrsim 2.65$. The Planck satellite is expected to be able to resolve f_{NL}^{equil} to $\mathcal{O}(10)$.

Having quantified the size of nongaussianity in axion inflation, we now turn our attention to its running. From eq. (3.28) we can see that

$$B_\zeta(k_1, k_2, k_3) = k^{-6} B_\zeta(1, x_2, x_3) \quad (4.12)$$

disregarding the mild, slow roll suppressed scale dependence of the vacuum solutions. The overall k^{-6} behavior reflects the near scale invariance of the bispectrum from axion inflation. Slight departures from scale invariance are quantified by the index $n_{NG} - 1 = \frac{d \ln |f_{NL}|}{d \ln k}$ which is easily seen to be proportional to the slow roll parameters ϵ, η , and hence negligible whenever the observational bounds on n_s are satisfied. We conclude that the running of nongaussianity is uninterestingly small in axion inflation.

4.3.2 The Shape of the Nongaussianity

In order to discuss the shape of the bispectrum, it is natural to extract the strong k^{-6}

scaling in (4.12) and define a “shape function” of the form

$$S(k_i) = N(k_1 k_2 k_3)^2 B_\zeta(k_i) \quad (4.13)$$

where the constant of proportionality, N , is arbitrary. This shape function coincides with the quantity that was plotted in many previous works, including [47] for example. For the case of interest, we have

$$S(\xi; x_2, x_3) \equiv \frac{1 + x_2^3 + x_3^3}{x_2 x_3} \frac{f_3(\xi; x_2, x_3)}{3 f_3(\xi; 1, 1)} \quad (4.14)$$

which is normalized so that $S(1, 1) = 1$. Note that the bispectrum is defined only in the region $x_2 + x_3 \geq 1$, which follows from the triangle inequality. Moreover, the bispectrum is symmetric under interchange of any two momenta, and therefore we can restrict to the region $x_3 \leq x_2 \leq 1$ to avoid considering the same configuration more than once.

We plot the shape function $S(x_2, x_3)$ from axion inflation in the left panel of Fig. 6. The bispectrum in this model depends on the parameter ξ . In practice, however, we find that only the size of the nongaussianity (quantified by f_{NL}^{equil}) depends strongly on ξ . The shape function $S(x_2, x_3)$, on the other hand, is very mildly dependent on ξ . In Fig. 6 we work in the $\xi \rightarrow \infty$ limit, in which case the shape becomes independent of model parameters. (This can be seen by using the large argument expansion (3.18) of \mathcal{I} in the expression (3.29) for f_3 .) For $\xi \sim \mathcal{O}(1)$ this figure would be nearly indistinguishable.

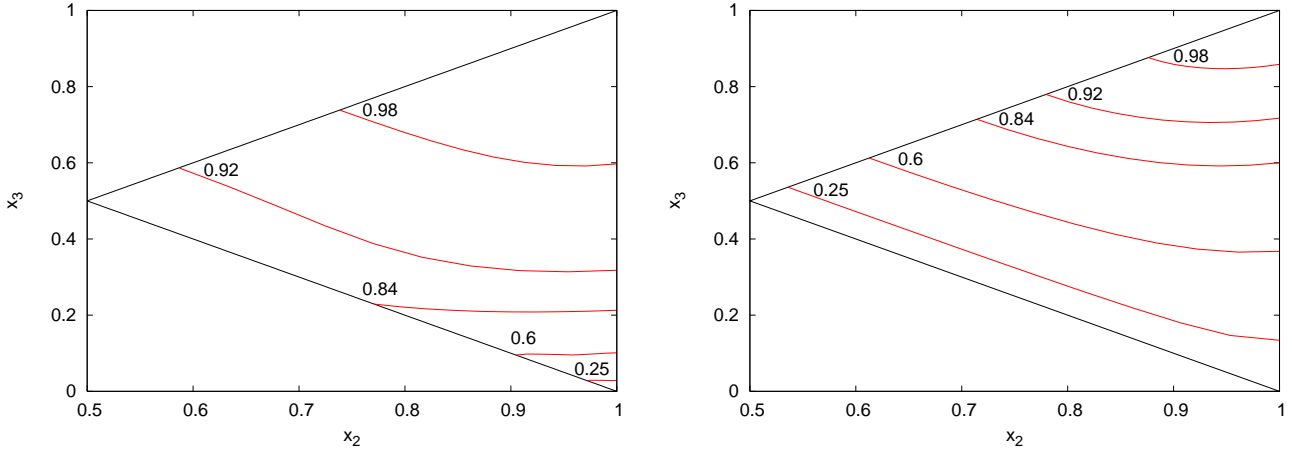


Figure 6: In the left panel we plot the shape function $S(x_2, x_3)$ in axion inflation, showing that this peaks on equilateral triangles. We work in the limit $\xi \rightarrow \infty$, however, this Figure would be nearly indistinguishable had we chosen $\xi = \mathcal{O}(1)$. In the right panel, for comparison, we plot the analogous shape function obtained from the standard equilateral template.

From Fig. 6 we see that the bispectrum from axion inflation peaks on equilateral triangles (corresponding to $x_2 = x_3 = 1$) and is thus qualitatively similar to the so-called equilateral template which is often employed to analyze CMB data [48, 39]

$$B_{\text{equil}}(k_i) \propto -\frac{1}{k_1^3 k_2^3} - \frac{1}{k_1^3 k_3^3} - \frac{1}{k_2^3 k_3^3} - \frac{2}{k_1^2 k_2^2 k_3^2} + \frac{1}{k_1 k_2^2 k_3^3} + (5 \text{ perms}) \quad (4.15)$$

(where the permutations only act on the last term). Equation (4.15) is the template that is used to obtain the WMAP7 limit on $f_{\text{NL}}^{\text{equil}} < 266$ which we employed in subsection 4.3.1. The shape function associated with the template (4.15) is plotted in the right panel of Fig. 6 for comparison with the analogous result from axion inflation.

To quantitatively compare the bispectrum from axion inflation to the equilateral template (4.15) we follow [47] and define a scalar product between any two bi-spectra as

$$B_1 \cdot B_2 \equiv \sum_{\vec{k}_i} B_1(\vec{k}_1, \vec{k}_2, \vec{k}_3) B_2(\vec{k}_1, \vec{k}_2, \vec{k}_3) / (\sigma_{k_1}^2 \sigma_{k_2}^2 \sigma_{k_3}^2) \quad (4.16)$$

where $\sigma_{k_i}^2$ is the variance of a given mode and the summation runs over all possible triangles. As shown in [47], this product is the best estimator for the overlapping of any two distributions: if we assume that the real data have the bi-spectrum B_1 , the template B_2 will produce a higher / lower value of nongaussianity according to how large / small the product (4.16) is [47]. To be quantitative, one defines the cosine of the “angle” between the two bi-spectra as [47]

$$\cos(B_1, B_2) \equiv \frac{B_1 \cdot B_2}{(B_1 \cdot B_1)^{1/2} (B_2 \cdot B_2)^{1/2}} \quad (4.17)$$

The cosine can be used to quickly estimate how well the limit given in the literature on some given template applies to a different shape [47].

We have computed the cosine of the “angle” between the bispectrum from axion inflation and the equilateral template for several different values of ξ . These results are reported in Table 1. There we see that the cosine depends only very weakly on ξ and saturates to a value $\cong 0.93$ in the limit $\xi \rightarrow \infty$. This confirms quantitatively our previous claim that the shape of nongaussianity is insensitive to ξ .

ξ	$\cos(B_{\text{inv.decay}}, B_{\text{equil}})$	$\cos(B_{\text{inv.decay}}, B_{\text{orth}})$
2	0.94	−0.093
3	0.94	−0.12
5	0.93	−0.13
∞	0.93	−0.15

Table 1: Cosine of the “Overlapping angle”, eq. (4.17), between the nongaussian shape generated by the inverse decay, and the equilateral (column 2) and orthogonal (column 3) templates, for different values of ξ .

Table 1 shows that $\cos(B_{\text{inv.decay}}, B_{\text{equil}})$ is very close to unity and hence we expect that the WMAP7 limit $-214 < f_{\text{NL}}^{\text{equil}} < 266$ can be applied also to axion inflation (at least to first approximation). This justifies our interpretation of the observational limit on nongaussianity in subsection 4.3.1.

Although the bispectrum from axion inflation is very similar to the equilateral template, the two shapes are not identical. It may be interesting to characterize the difference between these two shapes – indeed, this would become pressing in the event that Planck, or some other future mission, should detect a non-vanishing bispectrum on equilateral triangles. From Figure 6, we see that the bispectrum from inverse decay mostly differs from the

equilateral template for $x_2 \simeq x_3 \simeq 1/2$, corresponding to “flattened” triangles, where one side is twice the length of the other two. For such triangles, the bispectrum from inverse decay is significantly greater than the equilateral template. Hence, this provides a natural limit in which the two shapes may be distinguished.

The fact that axion inflation gives a large nongaussianity on flattened triangles implies a small but nontrivial overlap with the so-called “orthogonal template” that was introduced in [48]:

$$B_{\text{orth}}(k_i) \propto -\frac{3}{k_1^3 k_2^3} - \frac{3}{k_1^3 k_3^3} - \frac{3}{k_2^3 k_3^3} - \frac{8}{k_1^2 k_2^2 k_3^2} + \frac{3}{k_1 k_2^2 k_3^3} + (5 \text{ perms}) \quad (4.18)$$

(where the permutations only act on the last term). The corresponding shape function evaluates to +1 in the equilateral limit, and to -2 along the $x_3 = 1 - x_2$ boundary (which includes the flatten triangle configurations). In Table 1 we have computed the cosine of the angle between the bispectrum from axion inflation and the template (4.18). We find $\cos(B_{\text{inv.decay}}, B_{\text{orth}}) \cong -0.15$ (at large ξ , see Table 1 for intermediate ξ), while $\cos(B_{\text{equil}}, B_{\text{orth}}) \cong 0.21$. Therefore, the overlap with the orthogonal template may provide a useful tool to discriminate observationally between the nongaussianity from axion inflation and that of the equilateral template.

4.4 Large Field Inflation

Once the COBE normalization is imposed, the key phenomenological predictions of any inflationary model are the spectral index, n_s , the tensor-to-scalar ratio r , and the nonlinearity parameter f_{NL} . As we discussed previously, the spectral index in axion inflation has the standard form $n_s = 1 + 2\eta - 6\epsilon$ and requires no further discussion. The remaining observables depend on the coupling α/f , the axion velocity $\dot{\varphi}$ and the Hubble rate H . Out of these three quantities, we have defined the two combinations

$$\xi \equiv \frac{\alpha \dot{\varphi}}{2fH} \quad , \quad \mathcal{P}^{1/2} \equiv \frac{H^2}{2\pi|\dot{\varphi}|} \quad (4.19)$$

Both the two and three point correlation functions of ζ can be written solely in terms of these two combinations. Therefore both the power spectrum and the bi-spectrum are functions of ξ and \mathcal{P} only. The COBE normalization fixes $\mathcal{P}^{1/2}$ in terms of ξ , see Figure 4. Therefore, the predicted nonlinearity parameter f_{NL} is function of ξ only, see Figure 5.

The tensor to scalar ratio, on the other hand, is a function of a different combination of parameters, see eq. (4.4). For this reason, we cannot present it in a plot as a function of ξ only. This can only be achieved once the potential for the axion is specified, since this provides one additional (slow-roll) relation between the parameters of the model.

For a specific choice of inflationary potential $V(\phi)$, it is of interest to determine how the combinations (4.19), along with the tensor-to-scalar ratio, depend on parameters of the underlying theory. As we discuss in [18] and also in Section 6, in the models of axion inflation of interest, the axion/inflaton dynamics effectively occurs as in a large field inflationary model with potential:

$$V = \frac{\lambda_p}{p} \varphi^p \quad (4.20)$$

where λ_p has mass dimension $4 - p$. Using slow roll approximation, we find

$$H = \sqrt{\frac{\lambda_p}{3p}} \frac{(2pN)^{p/4}}{M_p^{1-p/2}}, \quad \dot{\phi} = \sqrt{\frac{p}{2N}} H M_p, \quad \xi = \frac{\alpha M_p}{2f} \sqrt{\frac{p}{2N}} \quad (4.21)$$

where N is the number of e-folds between the moment the CMB scales left the horizon and the end of inflation.

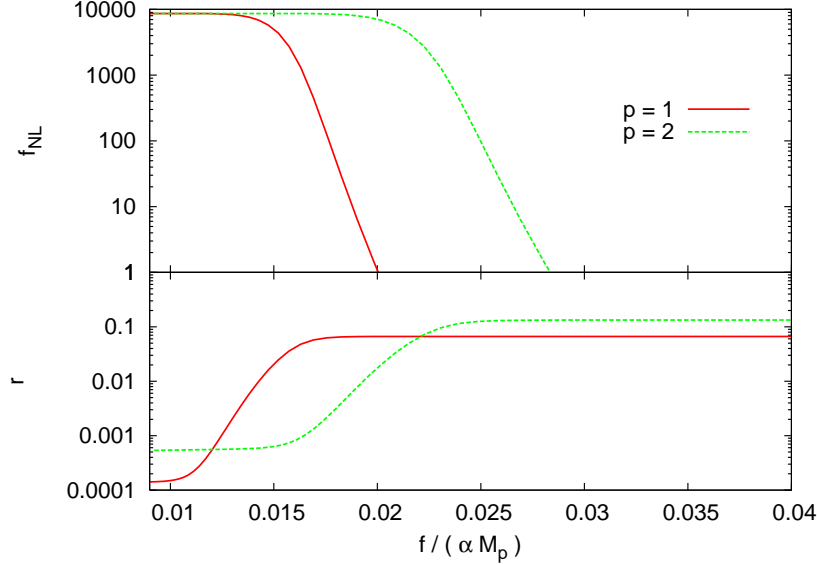


Figure 7: Predicted values for the equilateral f_{NL} parameter, and for the tensor to scalar ratio r in axion inflation models, as a function of the coupling of the axion to gauge fields, when the axion dynamics is effectively described by a potential $V \propto \varphi^p$, with $p = 1, 2$.

The value of λ_p is fixed by COBE normalization and we assume $N = 60$ e-foldings of inflation. Then, for any given value of p , all observational predictions can be written in terms of f/α only. In axion monodromy [12], $p = 1$; in most of the other models one expands the potential close to the minimum, where it is quadratic, $p = 2$. We therefore show in Figure 7 the predicted values of f_{NL} and of r for these two cases. We notice that, once f_{NL} is required to be below the WMAP7 bound, the standard value for r is recovered.

It is interesting to note that axion inflation models generically predict the *same* values of n_s and r as would be obtained in vanilla chaotic inflation. However, our scenario predicts also a large nongaussianity with a (nearly) equilateral shape. Axion inflation provides a rare example of a theory which predicts both a detectable tensor-to-scalar ratio *and* a large equilateral bispectrum. Note that, if such a nongaussian signal is eventually detected, then it will immediately fix the value of the coupling α/f . On the other hand, if Planck fails to detect nongaussianity then we will have a surprisingly stringent bound on the strongest axion-type couplings between the inflaton and *any* gauge field.

5. Cosmological Perturbation Theory

In section 2 we summarized the results of [18], providing a brief overview of the growth of gauge quanta and the production of inflaton fluctuations via inverse decay processes,

in the model (2.1). In section 3 we provided a detailed computation of the relevant correlation functions. There, and also in [18], we neglected scalar metric perturbations. It is intuitively clear that this should provide a sensible approximation when $f_{NL} \gtrsim 1$, since the nongaussianity associated with gravitational interactions is tiny as compared to the axion-type interactions that we study (1.3). In this section, we reconsider the analysis of section 3, this time consistently accounting for scalar metric perturbations. As would be expected, there will be no significant change in our key results.

5.1 ADM Formalism

We consider the theory (2.1), minimally coupled to Einstein gravity

$$S = \int d^4x \sqrt{-g} \left[\frac{M_p^2}{2} R - \frac{1}{2} \partial_\mu \varphi \partial^\mu \varphi - V(\varphi) - \frac{1}{4} F^{\mu\nu} F_{\mu\nu} - \frac{\alpha}{8f} \varphi \epsilon^{\mu\nu\alpha\beta} F_{\mu\nu} F_{\alpha\beta} \right] \quad (5.1)$$

The Levi-Civita tensor $\epsilon^{\mu\nu\alpha\beta}$ is defined as

$$\epsilon^{\mu\nu\alpha\beta} = \frac{1}{\sqrt{-g}} \eta^{\mu\nu\alpha\beta} \quad (5.2)$$

where the alternating symbol $\eta^{\mu\nu\alpha\beta}$ is +1 for even permutations of its indices, -1 for odd permutations, and zero otherwise.

To study the cosmological perturbations it is convenient to employ the Arnowitt-Deser-Misner (ADM) form of the metric

$$ds^2 = -N^2 dt^2 + h_{ij} (dx^i + N^i dt) (dx^j + N^j dt) \quad (5.3)$$

This parametrization has the advantage that the lapse function, N , and shift vector, N^i , appear in the action as Lagrange multipliers and hence are not dynamical degrees of freedom. We consider only scalar perturbations of the metric, and we employ the following gauge-fixing choices

$$\varphi(t, \mathbf{x}) = \phi(t) + \delta\varphi(t, \mathbf{x}) \quad (5.4)$$

$$h_{ij}(t, \mathbf{x}) = a^2(t) \delta_{ij} \quad (5.5)$$

$$A_0 = 0 \quad (5.6)$$

corresponding to Coulomb gauge for the vector, along with a flat slicing of the space-time (since the linearized solutions are used for A_μ , we can also impose $\partial_i A_i = 0$). It is now straightforward to vary the action (5.1) with respect to N , N^i , yielding the constraint equations. These constraints may be solved perturbatively so that N , N^i can be eliminated in favor of $\delta\varphi$, A_i . In this manner, we can derive an action which describes the dynamics of the scalar and gauge field fluctuations, and their leading interactions. This procedure has been described many times in the literature, following [21], and we do not report the explicit steps of the computation in full detail in this work. We also note that the cosmological perturbations in a model very similar to ours (differing only by the inclusion of the term $f(\varphi) F^2$ rather than $\varphi F \tilde{F}$) have already been studied in detail in [40].

5.2 Quadratic Action

The first non-trivial contribution in the perturbative expansion of (5.1) is the quadratic action for the fluctuations $\delta\varphi$, A_i , which give rise to the linearized equations of motion. To leading order in slow roll parameters, we find the following result for the quadratic terms involving the scalar field fluctuation

$$S_2^\varphi = \frac{1}{2} \int d\tau d^3x a^2 \left[(\partial_\tau \delta\varphi)^2 - \partial_i \delta\varphi \partial_i \delta\varphi + \left(a^2 m^2 - 3 \frac{\phi'^2}{M_p^2} \right) \delta\varphi^2 \right] \quad (5.7)$$

The dynamics of the gauge field, on the other hand, arise from the quadratic terms in the expansion of the Maxwell term in the action (5.1). We find

$$S_2^A = \frac{1}{2} \int d\tau d^3x \left[A'_i A'_i - \partial_j A_i \partial_j A_i + \frac{\alpha \phi'}{f} \epsilon_{ijk} A_i \partial_j A_k \right] \quad (5.8)$$

5.3 Cubic Action

The leading nongaussian effects are encoded by the cubic terms in the expansion of (5.1). These may be divided into two categories: interactions involving 3 scalars, and those involving one scalar and 2 vectors. In the first category, we find

$$S_3^{\varphi\varphi\varphi} = \int d\tau d^3x \frac{a^2 \phi'}{4\mathcal{H}M_p^2} \left[2\delta\varphi' \partial_i \partial^{-2}(\delta\varphi') \partial_i(\delta\varphi) - \delta\varphi(\delta\varphi')^2 - \delta\varphi \partial_i(\delta\varphi) \partial_i(\delta\varphi) \right] \quad (5.9)$$

where we have introduced the inverse Laplacian operator, defined through the relation $\partial^{-2} \partial_i \partial_i f = f$. As expected, this coincides with the well-known result for single field inflation [41, 42]. In (5.9) we have neglected a term proportional to V''' , which is usually sub-dominant in the slow roll expansion (see [43] for more discussion).

The interactions (5.9) would be present even in the absence of the gauge field A_μ . The contribution to the non-linearity parameter due to these terms is proportional to slow-roll parameters [21, 22, 43] and may therefore be neglected whenever $f_{NL} \gtrsim \mathcal{O}(1)$. In this work, we are therefore justified to ignore (5.9).

Much more interesting for our analysis is the second category of interaction: those involving one scalar $\delta\varphi$ and two gauge fields A_i . From the expansion of the Maxwell term F^2 in the action (5.1) we find

$$S_3^{\varphi AA} \supset \int d\tau d^3x \frac{\phi'}{\mathcal{H}M_p^2} \delta\varphi \left[-\frac{1}{4} A'_i A'_i - \frac{1}{8} F_{ij} F_{ij} + \frac{1}{2} \partial^{-2} \partial_\tau \partial_i (F_{ij} A'_j) \right] \quad (5.10)$$

which agrees with [40]. Finally, from the last term in (5.1) – the pseudo-scalar interaction – we have a contribution

$$S_3^{\varphi AA} \supset \int d\tau d^3x \frac{\alpha}{f} \delta\varphi \left[-\epsilon_{ijk} A'_i \partial_j A_k \right] \quad (5.11)$$

This interaction gives rise to the source term in (2.18) and was accounted for by the analysis of [18]. Equations (5.9,5.10,5.11) exhaust all interactions at cubic order.

The interaction (5.11) would be present even in the absence of metric fluctuations and was already accounted for in the analysis of [18] and section 3; this term simply arises from (1.3). On the other hand, the interactions (5.9) and (5.10) are “new” and arise due to the consistent inclusion of metric perturbations. It is easy to see already that these “new” interactions are completely negligible as compared to the pseudo-scalar coupling (5.11). To see this, note that the strength of the interaction (5.11) is controlled by the axion decay constant which is

$$\frac{\alpha}{f} \gtrsim 10^2 \frac{1}{M_p} \quad (5.12)$$

whenever the nongaussianity is observationally interesting [18]. The analogous coupling strength associated with the interaction (5.10) is instead

$$\frac{\phi'}{\mathcal{H}M_p^2} \sim \sqrt{\epsilon} \frac{1}{M_p} \sim 10^{-1} \frac{1}{M_p} \quad (5.13)$$

where in the last approximation we have assumed $n_s - 1 \sim \epsilon \sim 10^{-2}$, which is true in most axion inflation models. A similar line of reasoning can be used to verify that the trilinear interactions described by (5.9) are slow roll suppressed with respect to M_p^{-1} and are known to yield a negligible contribution to f_{NL} [43], as compared to (5.11).

This analysis indicates that the inclusion of metric perturbations has no significant impact in the interesting regime where $f_{NL} \gtrsim \mathcal{O}(1)$. This intuitive idea has been exploited on numerous occasions in the literature and was recently conjectured as a general decoupling principle in [44].

5.4 The Equations of Motion

In subsections 5.2 and 5.3 we derived the leading order contributions to the action for the scalar and gauge field fluctuations. In order to compute the nongaussianity in the model, there are two main approaches. We could use the in-in formalism to compute the bispectrum directly from the interactions (5.9,5.10,5.11). This is the approach that was adopted in [21, 22, 23, 40, 41]. Alternatively, we can use the perturbed action to derive the Klein-Gordon equation, and then compute the bispectrum following [43]. Both approaches give the same answer. In our case, we find the second approach to be more convenient, since it will allow us to take maximal advantage of our previous analysis [18].

Variation of the quadratic action (5.8) gives the linear equation of motion of the gauge field

$$\left[\partial_\tau^2 - \vec{\nabla}^2 - \frac{\alpha\phi'}{f} \vec{\nabla} \times \right] \vec{A}(\tau, \mathbf{x}) = 0 \quad (5.14)$$

This equation coincides exactly with our previous result (2.10). The consistent inclusion of metric perturbations has no impact on the linear fluctuations of the gauge field, which is a consequence of the fact that $\langle A_i \rangle = 0$.

Next, we consider the Klein-Gordon equation for the scalar fluctuation $\delta\varphi$. Variation of (5.7) along with (5.9,5.10,5.11) gives rise to the following equation of motion

$$\left[\partial_\tau^2 + 2\mathcal{H}\partial_\tau - \vec{\nabla}^2 + \left(a^2 m^2 - \frac{3\phi'^2}{M_p^2} \right) \right] \delta\varphi = \frac{\alpha}{f} a^2 \vec{E} \cdot \vec{B}$$

$$-\frac{a^2\phi'}{2M_p^2}\left[\frac{\vec{E}^2+\vec{B}^2}{2}+\frac{1}{a^4}\partial^{-2}\partial_\tau\left(a^4\vec{\nabla}\cdot\left(\vec{E}\times\vec{B}\right)\right)\right]+\dots \quad (5.15)$$

where the \dots on the last line denotes terms of order $(\delta\varphi)^2$ which arise due to variation of (5.9). As discussed in subsection 5.3, those terms contribute negligibly to the non-linearity parameter and can be ignored. As a check on our results, we have also derived equation (5.15) using a completely independent analysis: by expanding the Klein-Gordon equation to second order in perturbation theory and using the Einstein constraint equations to eliminate the metric fluctuations, following [45]. This analysis is briefly summarized in Appendix B.

Equation (5.15) differs from the result (2.18), that was derived by neglecting metric perturbations, in two respects. First off, equation (5.15) includes a slow-roll suppressed correction to the effective mass of the inflaton. The second discrepancy is the additional contribution to the source on the second line of (5.15), proportional to M_p^{-2} , arising from variation of (5.10). This contribution is parametrically of order $\sim \sqrt{\epsilon}\frac{\vec{E}^2+\vec{B}^2}{M_p}$ and may be neglected as compared to the term $\frac{\rho}{f}\vec{E}\cdot\vec{B}$, as discussed in subsection (5.3). The results of this subsection provide rigorous justification for our previous use of equation (2.18) to study the cosmological fluctuations in the model (5.1), in the $f_{NL} \gtrsim 1$ regime.

5.5 The Curvature Perturbation

It now remains to construct the curvature fluctuation on uniform density hypersurfaces, ζ . Beyond linear order, the relationship between ζ and $\delta\varphi$ is simplest to derive using the δN -formalism [46]. Because the gauge field does not contribute to the expansion history of the universe, we have the simple result on large scales:

$$\begin{aligned} \zeta \equiv \delta N &= \frac{\partial N}{\partial\varphi}\delta\varphi + \frac{1}{2}\frac{\partial^2 N}{\partial\varphi^2}(\delta\varphi)^2 + \dots \\ &= -\frac{H}{\dot{\phi}}\delta\varphi\left[1 + (2\epsilon - \eta)\frac{1}{2}\frac{H}{\dot{\phi}}\delta\varphi + \dots\right] \\ &\cong -\frac{H}{\dot{\phi}}\delta\varphi \end{aligned} \quad (5.16)$$

where on the last line we have neglected a slow-roll suppressed correction which contributes negligibly to f_{NL} in the case of interest.

Equations (5.15) and (5.16) are the main results of this section. Taken together, they provide a rigorous justification for the analysis of [18], and of the previous sections, which disregard the effect of metric perturbations.

6. Models of Axion Inflation

Our results concerning nongaussianity and inverse decay effects are quite general and may be applied directly to a variety axion inflationary models. Moreover, we expect that our

qualitative results may have also implications more broadly, for example in any multi-field scenario that involves a dynamical axion.⁸ In this section we survey some interesting microscopic constructions, both from field theory and string theory, and comment on the possibility of large nongaussianity. The key input which our analysis requires from a microscopic computation are the inflaton potential $V(\varphi)$, the decay constant f , and the dimensionless parameter α . Together, these determine the quantity

$$\xi \cong \frac{\alpha M_p^2}{2f} \left| \frac{V'}{V} \right| \quad (6.1)$$

that measures the strength of inverse decay effects. As we have seen, inverse decay processes cannot be neglected when $\xi \gtrsim \mathcal{O}(1)$ which roughly translated into $f/(\alpha M_p) \lesssim 10^{-2}$ for the most interesting models; see Figure 7. Whenever this inequality is satisfied, our findings strongly affect the phenomenology of these models.

In most of the scenarios that we survey, both $V(\varphi)$ and f tend to be fairly well understood while, on the other hand, the coefficient α is rather more model dependent (although calculable in principle). A detailed computation of α in each interesting scenario is beyond the scope of this paper and we leave such an analysis to future work. Although we generically expect $\alpha = \mathcal{O}(1)$, it should nevertheless be noted that this parameter contributes a source of theoretical uncertainty to what follows. We believe that our phenomenological results should provide a motivation for a detailed microscopic computation of α in the various scenarios discussed below.

6.1 Natural Inflation

The original natural inflation proposal [6] was based on the potential⁹

$$V(\varphi) \cong \Lambda^4 \left[1 - \cos \left(\frac{\varphi}{f} \right) \right] \quad (6.2)$$

For a spectral index $n_s \gtrsim 0.95$ this model requires $f \gtrsim \sqrt{8\pi} M_p$ [16], for which the inflaton dynamics shows little difference from that standard chaotic inflation with $V(\varphi) \cong \frac{1}{2}m^2\varphi^2$, at least for the present considerations. Such large values of f weaken the pseudo-scalar coupling $\varphi F\tilde{F}$ and inverse decay effects are negligible unless $\alpha \gtrsim \text{few} \times 10^2$. While specific situations have been constructed that can result in a large α - for instance in the some extra-dimensional model [15] - such large values conflicts with our general expectation that $\alpha = \mathcal{O}(1)$. We conclude that large nongaussianity seems unlikely in the simplest models of natural inflation.

In spite of apparent simplicity, however, the original natural inflation model [6] seems incompatible with UV completion. If we interpret φ as a PNCB then $f > M_p$ suggests

⁸This seems to be a generic expectation for closed string inflation models. In the effective SUGRA description the Kahler moduli τ_i are typically paired with axions θ_i into complex fields $T_i = \tau_i + i\theta_i$. Absent tuning, one expects the curvature of the scalar potential to be comparable in the τ_i and θ_i directions. Hence, nontrivial modular dynamics is typically also associated with nontrivial dynamics in the axion sector. See, for example, reference [49] for a review.

⁹See [50] for some proposed particle physics realizations. In [49] this model was realized in supergravity - the so-called “axion valley” model.

a global symmetry broken *above* the quantum gravity scale, where effective field theory is presumably not valid. Moreover, $f > M_p$ does not seem possible in a controlled limit of string theory [17] (which is the only known framework wherein such questions may be addressed). Hence, requiring the existence of a sensible UV completion automatically pushes us towards the regime $f \ll M_p$ where inverse decay effects are important. We will illustrate how this works in several explicit microscopic realizations below, however, it is clear that this trend applies more generally.

6.2 Double-Axion Inflation

Perhaps the simplest scenario to realize natural inflation with $f < M_p$ is the double-axion model proposed in [9]. This model is characterized by two axions, θ and ρ , whose potential

$$V(\theta, \rho) = \sum_{i=1}^2 \Lambda_i^4 \left[1 - \cos \left(\frac{\theta}{f_i} + \frac{\rho}{g_i} \right) \right] \quad (6.3)$$

arises from pseudo-scalar couplings to two different gauge groups: $\frac{\theta}{f_i} F_i \tilde{F}_i$ and $\frac{\rho}{g_i} F_i \tilde{F}_i$. For $f_1 g_2 = g_1 f_2$ one linear combination of θ and ρ becomes a flat direction of the potential, corresponding to an enhanced symmetry of the theory. Taking $f_1 g_2 \cong g_1 f_2$ the curvature in this direction becomes controllably flat and one may realize natural inflation with $f_i, g_i \ll M_p$. This scenario therefore provides a simple and compelling illustration of how the requirement of a UV completion leads to large nongaussianity in natural inflation.

6.3 N-flation

The N-flation model [10] is based on N axion fields φ_i , each with its own softly broken shift symmetry, resulting in a separable potential of the form

$$V(\varphi_i) \cong \sum_i \Lambda_i^4 \left[1 - \cos \left(\frac{\varphi_i}{f_i} \right) \right] \cong \sum_i \frac{1}{2} m_i^2 \varphi_i^2 \quad (6.4)$$

where, in the second equality, we have expanded in small field values $\varphi_i \ll f_i$. For $N \gg 1$ the assisted inflation mechanism [51] allows inflation to proceed even while all the decay constants are sub-Planckian, $f_i < M_p$. To a first approximation, the dynamics may be captured by a simple quadratic potential $V_{\text{eff}} \cong \frac{1}{2} m^2 \Phi^2$ for the collective field $\Phi^2 \equiv \sum_i \varphi_i^2$ [11]. Successful inflation requires that the collective field traverses a super-Planckian distance in field space, $\Delta \Phi \gtrsim M_p$. This is achievable with sub-Planckian φ_i provided the number of axions is sufficiently large. Roughly speaking we require [52]

$$N \sim 240 \left(\frac{M_p}{f_{\text{avg}}} \right)^2 \quad (6.5)$$

For $f_{\text{avg}} \sim 10^{-1} M_p$ we have $N \sim 10^3$ while $f_{\text{avg}} \sim 10^{-2} M_p$ would require $N \sim 10^6$.

N-flation makes sense as a purely field theoretical construction, however, much of the interest in this scenario arises because the exponentially large values of N that are required may be rather generic in string theory compactifications. In this context, axions arise on dimensional reduction, from integrating p -form gauge potentials over p -cycles (see [49, 53]

for a review). For example, in type IIB string theory one has axions b_i and c_i which arise, respectively, from integrating the Neveu-Schwarz (NS) and Ramond-Ramond (RR) 2-forms B_{MN} and C_{MN} over compact 2-cycles. Generically, instanton effects break the shift symmetry down to a subgroup $b_i \rightarrow b_i + (2\pi)f_i$, leading to periodic contributions to the effective potential (similarly for c_i).

The general framework described above shows how N-flation may arise within string theory: the low energy theory contains one axion for each independent cycle that B_{MN} can wrap and generic Calabi-Yau compactifications may contain exponentially large numbers of such cycles. There exist known examples with N as large as $\sim 10^5$ [54], however, to our knowledge there is no general theorem that prohibits finding Calabi-Yau manifolds with even larger values.

There may be many ways to realize N-flation in an explicit, stabilized string theory compactification. The first efforts in this direction were undertaken in [11] and further analyzed in [52]. See also [55] for an alternative construction.

In order to quantify inverse decay effects, we are most interested in the effective coupling between the collective field Φ and a given gauge field. The detailed derivation of such interactions is rather complicated and dependent on model building details (see [56, 57, 58] for a more detailed discussion). We will not attempt to estimate the effective value of α/f , but rather note that a generic spectrum is expected to contain some axions with $f_i \ll M_p$, hence strong inverse decay effects (and their associated nongaussianities) are at least plausible in N-flation.

6.4 Axion Monodromy Inflation

The axion monodromy model [12] is a string theoretic construction based on a single axion field. The key ingredient in this construction is a suitably wrapped brane which leads to a non-periodic contribution to the inflaton potential, explicitly breaking the shift symmetry. This monodromy in the moduli space allows the axion to develop a kinematically unbounded field range and accommodate the super-Planckian excursions required for large-field inflation. In the effective field theory description of this scenario, the inflaton potential has the form

$$V(\varphi) = \mu^3 \varphi + \Lambda^4 \cos\left(\frac{\varphi}{f}\right) \quad (6.6)$$

where the linear term arises from wrapped branes while the (subdominant) periodic modulation arises from instanton effects.

In [13] the decay constant for axion monodromy inflation was studied in detail. It was shown that microphysics bounds the allowed values as

$$0.06 \frac{g_s^{1/4}}{\mathcal{V}^{1/2}} < \frac{f}{M_p} < 0.9 g_s \quad (6.7)$$

where $g_s < 1$ is the string coupling and $\mathcal{V} \gg 1$ is the compactification volume in string units. This bound illustrates that $f \ll M_p$ in this model. If $\alpha = \mathcal{O}(1)$ can be realized in a consistent string compactification, then large nongaussianity is easily accommodated by axion monodromy inflation.

The small periodic modulation of the potential (6.6) is not important for determining the number of e -foldings of inflation, however, this term may nevertheless have an important role of the phenomenology of the model. In [13], and also in [59], it was shown that axion monodromy models can give rise to large resonant-type nongaussianities [27, 28, 60, 44]. Depending on the parameters, either resonant effects or inverse decay effects may dominate the bispectrum in the regime $f \ll M_p$. It would be interesting to study the combined observational impact of these two effects in future work.

6.5 Axion/4-Form Mixing

In [14] a realization of chaotic inflation was proposed which shares many features of the axion monodromy model discussed above. Here the axion φ mixes with a 4-form field strength through a term like $\varphi \epsilon^{\mu\nu\alpha\beta} F_{\mu\nu\alpha\beta}$. The theory also includes charged membranes which source a background for the conjugate momenta of the gauge field and break the axion shift symmetry, giving rise to a non-periodic potential $V(\varphi) \cong \frac{1}{2}m^2\varphi^2$ which is robust against a wide variety of corrections [61]. There is no obstruction to realizing inflation with $f \ll M_p$ in this model, and hence we generically expect that inverse decay processes may play an important role.

As in the case of axion monodromy, the power-law potential will generically be modulated by subdominant oscillatory features arising from instanton effects. In [61] nongaussianities were discussed. In general resonant effects may operate in concert with inverse decay processes, and it would be interesting to study their combined impact.

6.6 Dante’s Inferno

In [62] a model was proposed which consists of two axions, r and θ , with decay constants $f_r < f_\theta \ll M_p$. It is assumed that a linear combination of these receives a periodic potential from nonperturbative effects and, moreover, that some explicit shift symmetry breaking effect generates a non-periodic contribution $W(r)$. The effective potential therefore takes the form

$$V(r, \theta) = W(r) + \Lambda^4 \left[1 - \cos \left(\frac{r}{f_r} - \frac{\theta}{f_\theta} \right) \right] \quad (6.8)$$

For a power-law $W(r) = \mu^{4-p} r^p$ the dynamics of this model are well approximated by a single field with effective potential $V_{\text{eff}} = (f_r/f_\theta)^p \mu^{4-p} \varphi_{\text{eff}}^p$ and our analysis of inverse decay processes is directly applicable. The possibility of large nongaussianity follows immediately from $f_r < f_\theta \ll M_p$.

In [62] it was shown how to embed the model (6.8) within string theory as a modest extension of axion monodromy [12, 13]. It may be possible to realize such a potential from axion/4-form mixing, along the lines of [14, 61].

6.7 Multi-Field Scenarios

So far we have focused our attention on models which can, at least to first approximation, be described in terms of the dynamics of a single field (that may represent a collective excitation or be related non-trivially to the axions of the original theory). Such scenarios are appealing, because our analysis of inverse decay processes applies more-or-less without

modification. However, there are also a number of interesting inflation models which involve dynamical axions but are *not* well described in terms of a single dynamical field. (As discussed in footnote 8, this scenario seems especially natural in supergravity models.) In multi-field models with $f \ll M_p$ strong inverse decay effects are clearly possible, however, the detailed phenomenology is more complicated than what we have presented. Nevertheless, we expect that large nongaussianity should be possible. It would be interesting to explore multi-field inverse decay processes in future works.

There are a variety of interesting multi-field models involving axions, for example racetrack inflation [63, 64] and the axionic D3/D7 model [65]. Possibly, interesting models from the perspective of obtaining strong inverse decay effects, are based on the large volume compactification of [66, 67, 68]. For example, roulette inflation [69] – which generalizes [70] to incorporate the dynamics of the axionic partner of the Kahler modulus – is characterized by significant motion in the axion direction during the early stages of inflation. See [71, 72] for a discussion of axion/moduli couplings and decays in large volume inflationary scenarios. See also [73] for a related model.

7. Conclusions

Probably the greatest difficulty in inflationary model-building is to protect the required flatness of the potential from radiative corrections. One of the simplest solutions to this problem is to assume that the inflaton φ is a Pseudo-Nambu-Goldstone-Boson (PNGB). In this case the inflaton enjoys a shift symmetry $\varphi \rightarrow \varphi + \text{const}$, which is broken either explicitly or by quantum effects. In the limit of exact symmetry, the φ direction is flat and thus loop corrections to the potential are controlled by the smallness of the symmetry breaking. PNGBs, like the axion, are ubiquitous in particle physics: they arise whenever an approximate global symmetry is spontaneously broken and are plentiful in string compactifications. The idea of using a PNGB as inflaton was first put forward in [6]; in this minimal realization a single axion is present; it turns out that this model can produce a sufficiently flat spectrum of perturbations only if the axion decay constant f is above the Planck scale. This regime may be impossible to obtain in a controlled way: one can expect that gravity breaks the axionic shift symmetry at a smaller scale and it has also been conjectured that $f > M_p$ cannot be realized in string theory. These problems have been solved by a number of controlled realizations in which an axion, or a combination of axions, with an axion scale f a few orders of magnitude smaller than M_p , behaves effectively as a large field inflaton [9, 10, 11, 12, 13, 14, 15].

Axions are coupled to gauge fields through $\frac{\alpha}{f}\varphi F\tilde{F}$. More generally, from the perspective of effective field theory, a coupling $\varphi F\tilde{F}$ must be included whenever φ is pseudo-scalar. The dimensionless coupling α is a model-dependent quantity, however, from an field theoretical point of view, one does not expect it to be $\ll 1$. In this work we studied the phenomenological signatures induced by this coupling, keeping α/f as a free parameter. The crucial point of our work is that, in presence of this coupling, the motion of φ induces a tachyonic growth for one polarization δA of the gauge field. This amplification is most important for modes with physical wavelength comparable to the horizon $\lambda \sim 1/H$. These

produced gauge quanta inverse decay to produce inflaton fluctuations of a comparable wavelength, consistent with the general expectation from causality/locality. These inflaton modes then leave the horizon, and the corresponding density fluctuations, $\zeta_{\text{inv.decay}}$, become frozen and contribute to the observable cosmological perturbations. An analogous process leads to production of gravity waves. These scalar and tensor perturbations add up incoherently with those generated by the expansion of the universe (the standard “vacuum” modes, ζ_{vac} and h_{vac}).

We find that the amplitude of the perturbations generated by the inverse decay is an exponentially growing function of α/f . These modes dominate over the vacuum ones for $\alpha/f \gtrsim 10^{-2} M_p^{-1}$ (the precise value depending on the inflaton potential). Due to the exponential sensitivity on the coupling, there is only a small range in α/f for which both $\zeta_{\text{inv.decay}}$, and ζ_{vac} are relevant. For smaller values of α/f this new effect is completely negligible, and the standard results are recovered. For larger values, drastically new predictions are obtained. In particular, the main characteristic of $\zeta_{\text{inv.decay}}$ is that they are highly nongaussian: this is due to the fact that two gauge quanta participate in the inverse decay, and that the initial distribution of these quanta is itself gaussian (loosely speaking, $\zeta_{\text{inv.decay}}$ behaves as the square of a gaussian field, which is obviously not gaussian). As a consequence of this effect, the nongaussianity parameter f_{NL} also grows exponentially with α/f in the region of parameter space for which both ζ_{vac} and $\zeta_{\text{inv.decay}}$ are comparable. When ζ_{vac} can be neglected, f_{NL} saturates to about 8,600 in the equilateral configuration, well beyond the current WMAP limit $f_{NL}^{\text{equil}} < 266$. We find that the WMAP limit allows only values of α/f for which $\zeta_{\text{inv.decay}}$ can contribute $\lesssim 10\%$ to the power spectrum (see Figure 4). In this regime, also the spectrum of gravity waves produced by the inverse decay is much smaller than that from vacuum.

We have seen that nongaussianity from axion inflation is greatest for the equilateral configuration. This is related to causality/locality of the underlying inverse decay production mechanism. To understand this, intuitively, recall that a mode ζ_k is sourced by the inverse decay of gauge perturbations of comparable wavelength. Consequently, correlation

$$\langle \zeta_{k_1} \zeta_{k_3} \zeta_{k_3} \rangle \propto \int d^3 q_1 d^3 q_2 d^3 q_3 \langle \delta A_{q_1} \delta A_{k_1 - q_1} \delta A_{q_2} \delta A_{k_2 - q_2} \delta A_{q_3} \delta A_{k_3 - q_3} \rangle \quad (7.1)$$

is suppressed between modes of very different scale. We found that the shape of nongaussianity produced by this mechanism is very well reproduced by the equilateral template that has been widely studied in the literature. Quantitatively, the “cosine” between the two shapes (a measure of how much the two shapes coincide [47]) is $\cong 0.93$. Therefore, it is sensible to use the limits on f_{NL}^{equil} (obtained with the use of the equilateral template) to probe the current mechanism. On the other hand, the specific shape that we have computed is nevertheless distinct from the equilateral template and this fact might be useful to distinguish axion inflation from other models (for instance, the shape we found has a significant overlapping also with the orthogonal template). This issue can be more thoroughly explored when / if a nonzero value for f_{NL}^{equil} will be found in the data. The nongaussian signature of axion inflation, together with the requirement of standard results for the spectral tilt and the tensor-to-scalar ratio tensor of large field inflation, will allow

to falsify the mechanism (or, at least, the simplest version that we have studied here) in the near future. This is due to the fact that large field inflationary model provide much larger - and detectable - values of $n_s - 1$ and of r than many other classes of models.

We stress that the nongaussianity which we have studied is very different from the so called resonant nongaussianity [27, 28], that has been discussed previously [13, 59, 60, 44] in the context of axion monodromy inflation [12]. Axion monodromy is one of the particle physics realizations of axion inflation; in this model the periodic potential typical of axions is only a subdominant term in the inflaton potential, and it provides a nongaussian modulation of the inflaton perturbations. Depending on model parameters, either this effect or inverse decay processes may dominate the bispectrum.

It is remarkable that, in the mechanism we have studied, the large nongaussianity is obtained in a rather minimal way (simply by considering a coupling of the pseudo-scalar allowed by the symmetries of the model, and therefore expected in an effective field theory context). Nongaussianity is a measure of the strength of interactions of the inflaton, which are typically constrained to be small by the requirement of a flat potential. To obtain observable nongaussianity, previous studies have invoked non-standard field theories (involving small sound speed [23] or higher derivatives [5]), or initial conditions [23, 24, 25, 26], potentials with sharp features [27, 28], dissipative effects [3], fine-tuned inflationary trajectories [29] or post-inflationary effects (such as preheating [30, 31]). The mechanism we have studied circumvents the common lore result in a very novel way: the interaction that gives rise to large nongaussianity, eqn. (1.3), does not play any role in the background dynamics and is thus unconstrained by the requirement of slow roll. At the same time, the effect of this interaction (namely, the amplification of gauge quanta) persists during the entire inflationary phase, leading to a (nearly) scale invariant signature (as opposed to sporadic episodes of particle productions [32, 33, 34] that would lead to a highly localized nongaussianity in momentum space).

Finally, it is also remarkable that current observational limit on nongaussianity *already* place a surprisingly stringent bound on pseudo-scalar interactions of the form (1.3). As a comparison, the bound on the coupling of the much lighter QCD (or QCD-like) axion to photons is $\alpha_{\text{photons}}/f \lesssim \mathcal{O}(10^{-11}) \text{ GeV}^{-1}$, from energy loss in stars [74]. On the other hand, the mechanism that we have studied provides the bound $\alpha/f \lesssim \mathcal{O}(10^{-16}) \text{ GeV}^{-1}$ on the coupling of the pseudo-scalar inflaton with *any* gauge field. This provides a unique window for constraining – or perhaps probing – a large class of inflationary models. As we have reviewed in section 6, there are a variety of interesting multi-field models involving axions, many of which can be realized in string theory. While most of these studies provide the value of axion decay constant f and of the potential $V(\varphi)$, comparatively less attention has been devoted to compute the dimensionless coupling, α . We believe that our findings provide a strong motivation to undertake such a study.

Acknowledgments

We thank L. Sorbo for valuable discussions. This work was supported in part by DOE grant DE-FG02-94ER-40823 at UMN.

APPENDIX A: Gauge Field Mode Functions

In this appendix we discuss the solutions of equation (2.10), describing the unstable growth of gauge field fluctuations in the background of the slowly rolling inflaton (hence, ξ can be taken constant at leading order in the slow roll parameters). We require that the gauge field is initially in the adiabatic vacuum: $A_{\pm}(\tau, k) = e^{-ik\tau}/\sqrt{2k}$ for $k\tau \rightarrow -\infty$. The mathematical properties used here can be found in chapter 14 of [75]. The solution of (2.10) which satisfies this condition may be expressed in terms of Coulomb functions:

$$A_+(\tau, k) = \frac{1}{\sqrt{2k}} [G_0(\xi, -k\tau) + iF_0(\xi, -k\tau)] \quad (\text{A-1})$$

The production of gauge fluctuations is only interesting in the region of phase space $-k\tau \ll 2\xi$ and when $e^{\pi\xi} \gg 1$ (see the discussion in subsection 2.2). In this regime equation (A-1) may be very well approximated in terms of the modified Bessel function of the second kind, $K_\nu(z)$, as

$$A_+(\tau, k) \cong \sqrt{\frac{-2\tau}{\pi}} e^{\pi\xi} K_1 \left[2\sqrt{-2\xi k\tau} \right] \quad (\text{A-2})$$

The growth of the modes (A-2) saturates deep in the IR: for $-k\tau \rightarrow 0$ we have $A_+ \rightarrow e^{\pi\xi}/(2\sqrt{\pi k\xi})$ so that the physical electric and magnetic field vectors (2.7) decay sufficiently far outside the horizon. An inspection of the solutions shows that the interesting physical effects (for instance, the production of ζ^2 and ζ^3 correlators) take place in the region $(8\xi)^{-1} \lesssim -k\tau \lesssim 2\xi$ of phase space. In this regime we can take the large argument asymptotics of the Bessel function in (A-2) to obtain a very simple representation of the modes:

$$\begin{aligned} A_+(\tau, k) &\cong \frac{1}{\sqrt{2k}} \left(\frac{-k\tau}{2\xi} \right)^{1/4} e^{\pi\xi - 2\sqrt{-2\xi k\tau}} \\ A'_+(\tau, k) &\cong \sqrt{\frac{2k\xi}{-\tau}} A_+(\tau, k) \end{aligned} \quad (\text{A-3})$$

Throughout the majority of this paper we employ the representation (A-3) of the modes, for brevity of exposition. However, we have verified that none of our results changes significantly if we use the more accurate expression (A-2). Formally, the only effect of using (A-2) rather than (A-3) is that, for any two (rescaled) momenta \mathbf{q}_1 and \mathbf{q}_2 , the quantity

$$\mathcal{I} \left[c \left(|\mathbf{q}_1|^{1/2} + |\mathbf{q}_2|^{1/2} \right) \right] \simeq \int_0^\infty dx (\sin x - x \cos x) e^{-c(|\mathbf{q}_1|^{1/2} + |\mathbf{q}_2|^{1/2})\sqrt{x}} \quad (\text{A-4})$$

entering in eqs. (3.17) and (3.29) gets replaced by

$$\begin{aligned} \mathcal{I}_B &\equiv \frac{2c}{\pi} \frac{|\mathbf{q}_1|^{1/4} |\mathbf{q}_2|^{1/4}}{|\mathbf{q}_1|^{1/2} + |\mathbf{q}_2|^{1/2}} \int_0^\infty dx \sqrt{x} [\sin x - x \cos x] \\ &\times \left[|\mathbf{q}_1|^{1/2} K_1 \left(c |\mathbf{q}_1|^{1/2} \right) K_0 \left(c |\mathbf{q}_2|^{1/2} \right) + |\mathbf{q}_2|^{1/2} K_1 \left(c |\mathbf{q}_2|^{1/2} \right) K_0 \left(c |\mathbf{q}_1|^{1/2} \right) \right] \end{aligned} \quad (\text{A-5})$$

We verified that the results obtained with this replacement are in excellent agreement with those presented in the main text. Specifically, for the nongaussianity shape we obtained identical values to those presented in Table 1 for the nongaussianity shape; for the

nongaussianity parameter $f_{\text{NL}}^{\text{equil}}$ we obtained values consistent within a few percent with those shown in Figure 5. This explicitly confirms that the approximated expressions (A-3) are adequate for our analysis.

APPENDIX B: Perturbing the Klein-Gordon Equation

In section 5 we have derived a closed-form expression for Klein-Gordon equation (5.15) for the inflaton fluctuation $\delta\varphi$ in the theory (5.1). This was derived by employing the ADM decomposition of the metric (5.3) and integrating out the lapse N and shift N^i functions. However, one could arrive at precisely the same result by working directly with the equations of motion. Following [45], we expand the metric up to second order in perturbation theory as

$$g_{00} = -a^2(1 + 2\phi_1 + 2\phi_2) \quad (\text{B-1})$$

$$g_{0i} = a^2\partial_i(B_1 + B_2) \quad (\text{B-2})$$

$$g_{ij} = a^2\delta_{ij} \quad (\text{B-3})$$

Similar, we expand the scalar and gauge fields as

$$\varphi(t, \mathbf{x}) = \phi(t) + \delta_1\varphi(t, \mathbf{x}) + \delta_2\varphi(t, \mathbf{x}) \quad (\text{B-4})$$

$$A_\mu(t, \mathbf{x}) = (0, \delta_1 A_i(t, \mathbf{x}) + \delta_2 A_i(t, \mathbf{x})) \quad (\text{B-5})$$

In this gauge, the curvature perturbation on uniform density hypersurfaces is $\zeta = -\frac{H}{\phi}\delta\varphi$. The equation of motion for the scalar field is

$$\partial_\mu [\sqrt{-g}g^{\mu\nu}\partial_\nu\varphi] - \sqrt{-g}\frac{dV}{d\varphi} - \frac{\alpha}{8f}\eta^{\mu\nu\alpha\beta}F_{\mu\nu}F_{\alpha\beta} = 0 \quad (\text{B-6})$$

We expand (B-6) up to second order in perturbation theory, using the Einstein constraint equations to eliminate the metric fluctuations in order to close the system. At linear order we find

$$\left[\partial_\tau^2 + 2\mathcal{H}\partial_\tau - \vec{\nabla}^2 + \left(a^2m^2 - \frac{3\phi'^2}{M_p^2} \right) \right] \delta_1\varphi = 0 \quad (\text{B-7})$$

to leading order in slow roll. At second order in perturbation theory we find

$$\begin{aligned} & \left[\frac{\partial^2}{\partial_\tau^2} + 2\mathcal{H}\frac{\partial}{\partial_\tau} - \vec{\nabla}^2 + \left(a^2m^2 - \frac{3\phi'^2}{M_p^2} \right) \right] \delta_2\varphi = -\frac{\alpha}{f}\frac{1}{a^2}\epsilon_{ijk}\delta_1 A'_i\partial_j(\delta_1 A_k) \\ & + \frac{\phi'}{2a^2\mathcal{H}M_p^2} \left[-\frac{1}{2}\delta_1 A'_i\delta_1 A'_i - \frac{1}{4}\delta_1 F_{ij}\delta_1 F_{ij} + \partial^{-2}\partial_i(\delta_1 F_{ij}\delta_1 A'_j) \right] + \mathcal{O}[(\delta_1\varphi)^2] \end{aligned} \quad (\text{B-8})$$

to leading order in slow roll. These results coincide with those reported in section 5, which were derived starting from the action of the perturbations. Besides being a check on our algebra, this agreement strengthens the conclusions of [42], helping to establish the consistency of the two most popular approaches to nonlinear cosmological perturbation theory.

References

- [1] D. Baumann, A. Dymarsky, I. R. Klebanov, L. McAllister and P. J. Steinhardt, “A Delicate Universe,” *Phys. Rev. Lett.* **99**, 141601 (2007) [arXiv:0705.3837 [hep-th]].
- [2] A. Berera, “Warm Inflation,” *Phys. Rev. Lett.* **75**, 3218 (1995) [arXiv:astro-ph/9509049].
- [3] D. Green, B. Horn, L. Senatore and E. Silverstein, “Trapped Inflation,” *Phys. Rev. D* **80**, 063533 (2009) [arXiv:0902.1006 [hep-th]].
- [4] M. Alishahiha, E. Silverstein and D. Tong, “DBI in the sky,” *Phys. Rev. D* **70**, 123505 (2004) [arXiv:hep-th/0404084].
- [5] N. Barnaby, T. Biswas and J. M. Cline, “p-adic inflation,” *JHEP* **0704**, 056 (2007) [arXiv:hep-th/0612230]. N. Barnaby and J. M. Cline, “Large Nongaussianity from Nonlocal Inflation,” *JCAP* **0707**, 017 (2007) [arXiv:0704.3426 [hep-th]]. N. Barnaby and J. M. Cline, “Predictions for Nongaussianity from Nonlocal Inflation,” *JCAP* **0806**, 030 (2008) [arXiv:0802.3218 [hep-th]].
- [6] K. Freese, J. A. Frieman and A. V. Olinto, “Natural inflation with pseudo - Nambu-Goldstone bosons,” *Phys. Rev. Lett.* **65**, 3233 (1990).
- [7] F. C. Adams, J. R. Bond, K. Freese, J. A. Frieman and A. V. Olinto, “Natural Inflation: Particle Physics Models, Power Law Spectra For Large Scale Structure, And Constraints From Cobe,” *Phys. Rev. D* **47**, 426 (1993) [arXiv:hep-ph/9207245].
- [8] N. Arkani-Hamed, H. C. Cheng, P. Creminelli and L. Randall, “Extranatural inflation,” *Phys. Rev. Lett.* **90**, 221302 (2003) [arXiv:hep-th/0301218].
- [9] J. E. Kim, H. P. Nilles and M. Peloso, “Completing natural inflation,” *JCAP* **0501**, 005 (2005) [hep-ph/0409138].
- [10] S. Dimopoulos, S. Kachru, J. McGreevy and J. G. Wacker, “N-flation,” *JCAP* **0808**, 003 (2008) [hep-th/0507205].
- [11] R. Easther and L. McAllister, “Random matrices and the spectrum of N-flation,” *JCAP* **0605**, 018 (2006) [hep-th/0512102].
- [12] L. McAllister, E. Silverstein and A. Westphal, “Gravity Waves and Linear Inflation from Axion Monodromy,” *Phys. Rev. D* **82**, 046003 (2010) [arXiv:0808.0706].
- [13] R. Flauger, L. McAllister, E. Pajer, A. Westphal and G. Xu, “Oscillations in the CMB from Axion Monodromy Inflation,” *JCAP* **1006**, 009 (2010) [arXiv:0907.2916].
- [14] N. Kaloper and L. Sorbo, *Phys. Rev. Lett.* **102**, 121301 (2009) [arXiv:0811.1989].
- [15] M. M. Anber and L. Sorbo, “Naturally inflating on steep potentials through electromagnetic dissipation,” *Phys. Rev. D* **81**, 043534 (2010) [arXiv:0908.4089].
- [16] C. Savage, K. Freese and W. H. Kinney, “Natural Inflation: status after WMAP 3-year data,” *Phys. Rev. D* **74**, 123511 (2006) [hep-ph/0609144].
- [17] T. Banks, M. Dine, P. J. Fox and E. Gorbatov, “On the possibility of large axion decay constants,” *JCAP* **0306**, 001 (2003) [hep-th/0303252].
- [18] N. Barnaby and M. Peloso, “Large Nongaussianity in Axion Inflation,” arXiv:1011.1500 [hep-ph].

- [19] N. Barnaby, “Nongaussianity from Particle Production During Inflation,” *Adv. Astron.* **2010**, 156180 (2010) [arXiv:1010.5507].
- [20] V. Acquaviva, N. Bartolo, S. Matarrese and A. Riotto, “Second-order cosmological perturbations from inflation,” *Nucl. Phys. B* **667**, 119 (2003) [arXiv:astro-ph/0209156].
- [21] J. M. Maldacena, “Non-Gaussian features of primordial fluctuations in single field inflationary models,” *JHEP* **0305**, 013 (2003) [arXiv:astro-ph/0210603].
- [22] D. Seery and J. E. Lidsey, “Primordial non-gaussianities in single field inflation,” *JCAP* **0506**, 003 (2005) [arXiv:astro-ph/0503692].
- [23] X. Chen, M. x. Huang, S. Kachru and G. Shiu, “Observational signatures and non-Gaussianities of general single field inflation,” *JCAP* **0701**, 002 (2007) [arXiv:hep-th/0605045].
- [24] P. D. Meerburg, J. P. van der Schaar and P. S. Corasaniti, “Signatures of Initial State Modifications on Bispectrum Statistics,” *JCAP* **0905**, 018 (2009) [arXiv:0901.4044 [hep-th]].
- [25] P. D. Meerburg, J. P. van der Schaar and M. G. Jackson, “Bispectrum signatures of a modified vacuum in single field inflation with a small speed of sound,” arXiv:0910.4986 [hep-th].
- [26] R. Holman and A. J. Tolley, “Enhanced Non-Gaussianity from Excited Initial States,” *JCAP* **0805**, 001 (2008) [arXiv:0710.1302 [hep-th]].
- [27] X. Chen, R. Easther and E. A. Lim, “Large non-Gaussianities in single field inflation,” *JCAP* **0706**, 023 (2007) [arXiv:astro-ph/0611645].
- [28] X. Chen, R. Easther and E. A. Lim, “Generation and Characterization of Large Non-Gaussianities in Single Field Inflation,” *JCAP* **0804**, 010 (2008) [arXiv:0801.3295 [astro-ph]].
- [29] G. I. Rigopoulos, E. P. S. Shellard and B. J. W. van Tent, “Large non-Gaussianity in multiple-field inflation,” *Phys. Rev. D* **73**, 083522 (2006) [arXiv:astro-ph/0506704].
F. Vernizzi and D. Wands, “Non-Gaussianities in two-field inflation,” *JCAP* **0605**, 019 (2006) [arXiv:astro-ph/0603799].
C. T. Byrnes, K. Y. Choi and L. M. H. Hall, “Conditions for large non-Gaussianity in two-field slow-roll inflation,” *JCAP* **0810**, 008 (2008) [arXiv:0807.1101 [astro-ph]].
C. T. Byrnes and G. Tasinato, “Non-Gaussianity beyond slow roll in multi-field inflation,” *JCAP* **0908**, 016 (2009) [arXiv:0906.0767 [astro-ph.CO]].
X. Chen and Y. Wang, “Quasi-Single Field Inflation and Non-Gaussianities,” arXiv:0911.3380 [hep-th].
- [30] N. Barnaby and J. M. Cline, “Nongaussian and nonscale-invariant perturbations from tachyonic preheating in hybrid inflation,” *Phys. Rev. D* **73**, 106012 (2006) [arXiv:astro-ph/0601481].
N. Barnaby and J. M. Cline, “Nongaussianity from Tachyonic Preheating in Hybrid Inflation,” *Phys. Rev. D* **75**, 086004 (2007) [arXiv:astro-ph/0611750].
- [31] J. R. Bond, A. V. Frolov, Z. Huang and L. Kofman, “Non-Gaussian Spikes from Chaotic Billiards in Inflation Preheating,” *Phys. Rev. Lett.* **103**, 071301 (2009) [arXiv:0903.3407 [astro-ph.CO]].

- [32] N. Barnaby, Z. Huang, L. Kofman and D. Pogosyan, “Cosmological Fluctuations from Infra-Red Cascading During Inflation,” *Phys. Rev. D* **80**, 043501 (2009) [arXiv:0902.0615 [hep-th]].
- [33] N. Barnaby and Z. Huang, “Particle Production During Inflation: Observational Constraints and Signatures,” *Phys. Rev. D* **80**, 126018 (2009) [arXiv:0909.0751 [astro-ph.CO]].
- [34] N. Barnaby, “On Features and Nongaussianity from Inflationary Particle Production,” arXiv:1006.4615 [astro-ph.CO].
- [35] M. LoVerde, A. Miller, S. Shandera *et al.*, “Effects of Scale-Dependent Non-Gaussianity on Cosmological Structures,” *JCAP* **0804**, 014 (2008). [arXiv:0711.4126 [astro-ph]].
- [36] L. Sorbo, “Parity violation in the Cosmic Microwave Background from a pseudoscalar inflaton,” arXiv:1101.1525 [astro-ph.CO].
- [37] S. Saito, K. Ichiki, A. Taruya, “Probing polarization states of primordial gravitational waves with CMB anisotropies,” *JCAP* **0709**, 002 (2007). [arXiv:0705.3701 [astro-ph]].
- [38] V. Gluscevic, M. Kamionkowski, “Testing Parity-Violating Mechanisms with Cosmic Microwave Background Experiments,” *Phys. Rev. D* **81**, 123529 (2010). [arXiv:1002.1308 [astro-ph.CO]].
- [39] E. Komatsu *et al.* [WMAP Collaboration], “Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation,” arXiv:1001.4538 [astro-ph.CO].
- [40] D. Seery, “Magnetogenesis and the primordial non-gaussianity,” *JCAP* **0908**, 018 (2009) [arXiv:0810.1617 [astro-ph]].
- [41] D. Seery and J. E. Lidsey, “Primordial non-gaussianities from multiple-field inflation,” *JCAP* **0509**, 011 (2005) [arXiv:astro-ph/0506056].
- [42] K. A. Malik, D. Seery and K. N. Ananda, “Different approaches to the second order Klein-Gordon equation,” *Class. Quant. Grav.* **25**, 175008 (2008) [arXiv:0712.1787 [astro-ph]].
- [43] D. Seery, K. A. Malik and D. H. Lyth, “Non-gaussianity of inflationary field perturbations from the field equation,” *JCAP* **0803**, 014 (2008) [arXiv:0802.0588 [astro-ph]].
- [44] L. Leblond and E. Pajer, “Resonant Trispectrum and a Dozen More Primordial N-point functions,” arXiv:1010.4565 [hep-th].
- [45] K. A. Malik, “A not so short note on the Klein-Gordon equation at second order,” *JCAP* **0703**, 004 (2007) [arXiv:astro-ph/0610864].
- [46] D. H. Lyth, K. A. Malik and M. Sasaki, “A general proof of the conservation of the curvature perturbation,” *JCAP* **0505**, 004 (2005) [arXiv:astro-ph/0411220].
- [47] D. Babich, P. Creminelli and M. Zaldarriaga, “The shape of non-Gaussianities,” *JCAP* **0408**, 009 (2004) [arXiv:astro-ph/0405356].
- [48] L. Senatore, K. M. Smith and M. Zaldarriaga, “Non-Gaussianities in Single Field Inflation and their Optimal Limits from the WMAP 5-year Data,” *JCAP* **1001**, 028 (2010) [arXiv:0905.3746 [astro-ph.CO]].
- [49] R. Kallosh, “On Inflation in String Theory,” *Lect. Notes Phys.* **738**, 119 (2008) [arXiv:hep-th/0702059].

- [50] F. C. Adams, J. R. Bond, K. Freese *et al.*, “Natural inflation: Particle physics models, power law spectra for large scale structure, and constraints from COBE,” *Phys. Rev.* **D47**, 426-455 (1993). [hep-ph/9207245].
- [51] A. R. Liddle, A. Mazumdar and F. E. Schunck, “Assisted inflation,” *Phys. Rev. D* **58**, 061301 (1998) [arXiv:astro-ph/9804177].
- [52] R. Kallosh, N. Sivanandam and M. Soroush, “Axion Inflation and Gravity Waves in String Theory,” *Phys. Rev. D* **77**, 043501 (2008) [arXiv:0710.3429 [hep-th]].
- [53] D. Baumann and L. McAllister, “Advances in Inflation in String Theory,” *Ann. Rev. Nucl. Part. Sci.* **59**, 67 (2009) [arXiv:0901.0265 [hep-th]].
- [54] P. Candelas, E. Perevalov and G. Rajesh, “Toric geometry and enhanced gauge symmetry of F-theory/heterotic vacua,” *Nucl. Phys. B* **507**, 445 (1997) [arXiv:hep-th/9704097].
- [55] T. W. Grimm, “Axion inflation in type II string theory,” *Phys. Rev.* **D77**, 126007 (2008). [arXiv:0710.3883 [hep-th]].
- [56] D. R. Green, “Reheating Closed String Inflation,” *Phys. Rev. D* **76**, 103504 (2007) [arXiv:0707.3832 [hep-th]].
- [57] J. Braden, L. Kofman and N. Barnaby, “Reheating the Universe After Multi-Field Inflation,” *JCAP* **1007**, 016 (2010) [arXiv:1005.2196 [hep-th]].
- [58] M. M. Anber and L. Sorbo, “N-flationary magnetic fields,” *JCAP* **0610**, 018 (2006) [arXiv:astro-ph/0606534].
- [59] S. Hannestad, T. Haugbolle, P. R. Jarnhus and M. S. Sloth, “Non-Gaussianity from Axion Monodromy Inflation,” *JCAP* **1006**, 001 (2010) [arXiv:0912.3527 [hep-ph]].
- [60] R. Flauger and E. Pajer, “Resonant Non-Gaussianity,” arXiv:1002.0833 [hep-th].
- [61] N. Kaloper, A. Lawrence and L. Sorbo, “An Ignoble Approach to Large Field Inflation,” arXiv:1101.0026 [hep-th].
- [62] M. Berg, E. Pajer and S. Sjors, “Dante’s Inferno,” *Phys. Rev. D* **81**, 103535 (2010) [arXiv:0912.1341 [hep-th]].
- [63] J. J. Blanco-Pillado *et al.*, “Racetrack inflation,” *JHEP* **0411**, 063 (2004) [arXiv:hep-th/0406230].
- [64] J. J. Blanco-Pillado *et al.*, “Inflating in a better racetrack,” *JHEP* **0609**, 002 (2006) [arXiv:hep-th/0603129].
- [65] C. P. Burgess, J. M. Cline, M. Postma, “Axionic D3-D7 Inflation,” *JHEP* **0903**, 058 (2009). [arXiv:0811.1503 [hep-th]].
- [66] V. Balasubramanian and P. Berglund, “Stringy corrections to Kahler potentials, SUSY breaking, and the cosmological constant problem,” *JHEP* **0411**, 085 (2004) [arXiv:hep-th/0408054].
- [67] V. Balasubramanian, P. Berglund, J. P. Conlon and F. Quevedo, “Systematics of Moduli Stabilisation in Calabi-Yau Flux Compactifications,” *JHEP* **0503**, 007 (2005) [arXiv:hep-th/0502058].
- [68] J. P. Conlon, F. Quevedo and K. Suruliz, “Large-volume flux compactifications: Moduli spectrum and D3/D7 soft supersymmetry breaking,” *JHEP* **0508**, 007 (2005) [arXiv:hep-th/0505076].

- [69] J. R. Bond, L. Kofman, S. Prokushkin and P. M. Vaudrevange, “Roulette inflation with Kaehler moduli and their axions,” *Phys. Rev. D* **75**, 123511 (2007) [arXiv:hep-th/0612197].
- [70] J. P. Conlon and F. Quevedo, “Kaehler moduli inflation,” *JHEP* **0601**, 146 (2006) [arXiv:hep-th/0509012].
- [71] J. P. Conlon and F. Quevedo, “Astrophysical and Cosmological Implications of Large Volume String Compactifications,” *JCAP* **0708**, 019 (2007) [arXiv:0705.3460 [hep-ph]].
- [72] N. Barnaby, J. R. Bond, Z. Huang and L. Kofman, “Preheating After Modular Inflation,” *JCAP* **0912**, 021 (2009) [arXiv:0909.0503 [hep-th]].
- [73] A. Misra, P. Shukla, “Large Volume Axionic Swiss-Cheese Inflation,” *Nucl. Phys.* **B800**, 384-400 (2008). [arXiv:0712.1260 [hep-th]].
- [74] K. Nakamura et al. (Particle Data Group), *J. Phys. G* **37**, 075021 (2010).
- [75] M. Abramowitz and I. A. Stegun , “Handbook of Mathematical Functions”